GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:00:02 ; Search time 3180 Seconds
(without alignments)
6998.376 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTACAAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	544	100.0	544	6	AR259035	AR259035 Sequence
2	507.2	93.2	733	5	AF176420	AF176420 Deinagkis
3	458.2	84.2	656	5	AB036880	AB036880 Deinagkis
4	377.4	69.4	497	5	AY091758	AY091758 Deinagkis
5	376.2	69.2	678	5	TFLFIXA	D83331 Trimeresuru
6	354.8	65.2	683	5	AB046491	AB046491 Trimeresu
7	335.8	61.7	661	5	AF540646	AF540646 Deinagkis
8	313.4	57.6	630	5	AF190827	AF190827 Gloydius
9	312	57.4	666	5	AF540645	AF540645 Deinagkis
10	306.2	56.3	721	6	AX427204	AX427204 Sequence
11	306	56.2	704	5	AF102901	AF102901 Deinagkis
12	297.4	54.7	681	5	AY149341	AY149341 Trimeresu
13	288.2	53.0	800	5	CDRNACVXA	Y16348 Crotalus du
14	288	52.9	725	6	AX427210	AX427210 Sequence
15	285.8	52.5	580	6	AX427207	AX427207 Sequence
16	276.2	50.8	632	5	AF125309	AF125309 Gloydius
17	248	45.6	508	5	AF463522	AF463522 Deinagkis
18	241.2	44.3	679	5	AB019615	AB019615 Agkistrod
19	226.4	41.6	690	6	AR026653	AR026653 Sequence
20	226.4	41.6	690	6	AX404807	AX404807 Sequence
21	226.4	41.6	690	6	BD143687	BD143687 Protein h
22	222.4	40.9	495	5	AY091762	AY091762 Deinagkis
23	218	40.1	508	5	AY091759	AY091759 Deinagkis
24	213.8	39.3	477	5	AY099321	AY099321 Protoboth
25	212	39.0	688	5	AF540647	AF540647 Deinagkis
26	211.2	38.8	408	5	AF244900	AF244900 Callosela
27	186.6	34.3	456	6	AX427206	AX427206 Sequence
28	182.8	33.6	738	5	CDRNACVXB	Y16349 Crotalus du
29	181.8	33.4	698	5	TFLFIXB	D83332 Trimeresuru
30	175	32.2	432	6	AX427209	AX427209 Sequence
31	167.6	30.8	685	5	AF102902	AF102902 Deinagkis
32	167	30.7	681	5	AY149340	AY149340 Trimeresu
33	166	30.5	474	6	AX427212	AX427212 Sequence
34	164.4	30.2	574	5	AF350324	AF350324 Deinagkis
35	154.4	28.4	561	5	AY268947	AY268947 Echis car
36	145	26.7	592	5	AB036881	AB036881 Deinagkis
37	144	26.5	634	5	AF197915	AF197915 Gloydius
38	133.2	24.5	664	5	AB019616	AB019616 Agkistrod
39	128.2	23.6	619	5	AF540648	AF540648 Deinagkis
40	104.4	19.2	478	5	AY091761	AY091761 Deinagkis
41	102.6	18.9	272	6	AR026652	AR026652 Sequence
42	101.6	18.7	478	5	AY091756	AY091756 Deinagkis
43	98.8	18.2	565	5	AY268948	AY268948 Echis car
44	98.8	18.2	583	5	AF125310	AF125310 Gloydius
45	93.6	17.2	478	5	AY091755	AY091755 Deinagkis

ALIGNMENTS

RESULT 1
AR259035
LOCUS AR259035 544 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6489451.
ACCESSION AR259035
VERSION AR259035.1 GI:27309520
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 544)
AUTHORS Li,B.X. and Cheng,X.
TITLE Antithrombosis enzyme from the snake venom of agkistrodon acutus
JOURNAL Patent: US 6489451-A 1 03-DEC-2002;
FEATURES Location/Qualifiers

```
     source          1. .544
                     /organism="unknown"
BASE COUNT       158 a    122 c    143 g    121 t
ORIGIN
```

Query Match 100.0%; Score 544; DB 6; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.2e-151;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540

Qy        541 AAAA 544
              ||||
Db        541 AAAA 544
```

```
RESULT 2
AF176420
LOCUS       AF176420                733 bp    mRNA    linear   VRT 30-APR-2002
DEFINITION  Deinagkistrodon acutus clone 2100490 agkisacutacin A chain mRNA,
            complete cds.
ACCESSION   AF176420 AY091757
VERSION     AF176420.2  GI:20273043
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J. and Liu,J.
  TITLE     cDNA sequencing and analysis of eleven C-type lectin-like protein
            subunits from Agkistrodon acutus
  JOURNAL   Sheng wu hua hsueh yu sheng wu wu li hsueh pao (2002) In press
REFERENCE   2  (bases 1 to 733)
  AUTHORS   Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and
            Liu,J.
  TITLE     Agkisacutacin, a new fibrinolytic & anti-platelet protein from
            Agkistrodon acutus venom
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     A chain of agkisacutacin from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 733)
  AUTHORS   Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
REFERENCE   5  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
  REMARK    Sequence update by submitter
COMMENT     On Apr 23, 2002 this sequence version replaced gi:6715112.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100490"
                     /tissue_type="venom"
     CDS             89. .547
                     /note="consists of two heterologous subunits (A and B
                     chains) linked by an inter-subunit disulfide bond; c-type
                     lectin family member"
                     /codon_start=1
                     /product="agkisacutacin A chain"
                     /protein_id="AAF26286.2"
                     /db_xref="GI:20273044"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCSSGWSSYEGHCYKVFKQSK
                     TWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHIETGFHKWENFYCEQQDPFVCEA"
BASE COUNT       191 a    173 c    204 g    165 t
ORIGIN
```

Query Match 93.2%; Score 507.2; DB 5; Length 733;
Best Local Similarity 98.0%; Pred. No. 1.3e-140;
Matches 535; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 244

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 304

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 364

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 424

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 484

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 544

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        545 TAGTCTGAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 604

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
```

```
Db        605 CCAACCCTGCCT-GCCACAATCTCTGCTCTGCACCCTTCCCTCAACGGATGCTGCGTGTA 663

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGCCT 538

Db        664 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGGTCCAATAAATTCTGCCTAGCCT 723

Qy        539 GAAAAA 544

Db        724 GAAAAA 729
```

RESULT 3
AB036880
LOCUS AB036880 656 bp mRNA linear VRT 23-MAY-2002
DEFINITION Deinagkistrodon acutus acp-a mRNA for anticoagulant protein A, complete cds.
ACCESSION AB036880
VERSION AB036880.1 GI:8980618
KEYWORDS anticoagulant protein A.
SOURCE Deinagkistrodon acutus
ORGANISM Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE 1
AUTHORS Tani,A., Ogawa,T., Nose,T., Nikandrov,N.N., Deshimaru,M., Chijiwa,T., Chang,C.C., Fukumaki,Y. and Ohno,M.
TITLE Characterization, primary structure and molecular evolution of anticoagulant protein from Agkistrodon actus venom
JOURNAL Toxicon 40 (6), 803-813 (2002)
MEDLINE 21975857
REFERENCE 2 (bases 1 to 656)
AUTHORS Ogawa,T. and Tani,A.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate School of Agricultural Science; 1-1, Tsutusmidori Amamiyamachi, Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp, URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808, Fax:81-22-717-8807)
FEATURES Location/Qualifiers
source 1. .656
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/tissue_type="venom gland"
gene 1. .656
/gene="acp-a"
CDS 50. .508
/gene="acp-a"
/codon_start=1
/product="anticoagulant protein A"
/protein_id="BAA99281.1"
/db_xref="GI:8980619"
/translation="MGRFIFVSFGLLVVYLSLSGTAADCSSSWSSYEGHCYKAFKQSK
TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
sig_peptide 50. .118
/gene="acp-a"
BASE COUNT 173 a 148 c 185 g 150 t
ORIGIN

Query Match 84.2%; Score 458.2; DB 5; Length 656;
Best Local Similarity 94.4%; Pred. No. 6.3e-126;
Matches 488; Conservative 0; Mismatches 23; Indels 6; Gaps 1;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db        146 GAAGGGCATTGCTATAAGGCCTTCAAACAATCTAAGACCTGGGCAGATGCAGAGAGCTTC 205

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        206 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 265

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        266 TTTGTGGCCCATCTGATTGCTCAGAAGATAAAGTCGGCCAAAATCCATGTCTGGATCGGA 325

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        326 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 385

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        386 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACAAAGCGACA 445

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Db        446 GGGTTTCGTAAGTGGGAGAATTTTTACTGTGAACAACGAGATCCATTTGTATGCGAAGCA 505

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420

Db        506 TAGTCTGAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGAAAGTCCCCCACCCCAGCCC 565

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480

Db        566 CCATCC------TGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 619

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG 517

Db        620 GCTGGGTCTGGTGTTGCTACTCCTGATGGGCCAGAAG 656
```

RESULT 4
AY091758
LOCUS AY091758 497 bp mRNA linear VRT 13-MAY-2002
DEFINITION Deinagkistrodon acutus clone 2100590 ACF 1/2 A-chain mRNA, complete cds.
ACCESSION AY091758
VERSION AY091758.1 GI:20562936
KEYWORDS .
SOURCE Deinagkistrodon acutus
ORGANISM Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE 1 (bases 1 to 497)
AUTHORS Yu,H., Xiang,K., Wang,Y. and Liu,J.
TITLE A chain of ACF 1/2 from Deinagkistrodon acutus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 497)
AUTHORS Yu,H., Xiang,K., Wang,Y. and Liu,J.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular Biology, University of Science and Technology of China, School of Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES Location/Qualifiers
source 1. .497
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"
/db_xref="taxon:36307"
/clone="2100590"
CDS 1. .459
/note="C-type lectin family member"
/codon_start=1
/product="ACF 1/2 A-chain"
/protein_id="AAM22786.1"
/db_xref="GI:20562937"
/translation="MGRFIFVSFGLLVVFLSLSGTAADCSSSWSSYEGHCYKAFKQSK
TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
BASE COUNT 135 a 102 c 143 g 117 t
ORIGIN

Query Match 69.4%; Score 377.4; DB 5; Length 497;
Best Local Similarity 97.2%; Pred. No. 9.9e-102;

```
  Matches  384;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||| |||||||||||||||||||| |||||||||||||||||||
Db         97 GAAGGGCATTGCTACAAGGCCTTCAAACAATCTAAGACCTGGGCAGATGCAGAGAGCTTC 156

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 216

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 TTTGTGGCCCATCTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 276

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 336

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||||||||||||||||||||||||||||||||||||||||| || ||||||||
Db        337 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACAAAGCGACA 396

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||| |||||||||||||||||||||||||||| |||||||| ||||||||||||
Db        397 GGGTTTCGTAAGTGGGAGAATTTTTACTGTGAACAACGAGATCCTTTCGTCTGCGAGGCA 456

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAA 395
              ||||||||||||||||||| ||||||||||||||
Db        457 TAGTCTGAAGATCCAGCTGAGTGAAGTCTGGAGAA 491


RESULT 5
TFLFIXA
LOCUS       TFLFIXA                   678 bp    mRNA    linear   VRT 06-FEB-1999
DEFINITION  Trimeresurus flavoviridis mRNA for factor IX/factor X binding
            protein A chain, complete cds.
ACCESSION   D83331
VERSION     D83331.1  GI:1402639
KEYWORDS    facotr IX/factor X binding protein A chain.
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Matsuzaki,R., Yoshiara,E., Yamada,M., Shima,K., Atoda,H. and
            Morita,T.
  TITLE     cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
            protein from snake venom
  JOURNAL   Biochem. Biophys. Res. Commun. 220 (2), 382-387 (1996)
  MEDLINE   96184662
   PUBMED   8645314
REFERENCE   2  (bases 1 to 678)
  AUTHORS   Morita,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1996) Takashi Morita, Meiji College of Pharmacy,
            Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
            (Tel:0424-21-0101(ex.429), Fax:0424-21-1489)
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
     5'UTR           1. .49
     CDS             50. .508
                     /codon_start=1
                     /product="factor IX/factor X binding protein A chain"
                     /protein_id="BAA11887.1"
                     /db_xref="GI:1402640"
                     /translation="MGRFIFMSFGLLVVAASLRGTGADCLSGWSSYEGHCYKAFEKYK
                     TWEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIWIGLRVQGKVKQC
                     NSEWSDGSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCGQQNPFVCEA"
     3'UTR           509. .678

     polyA_signal    659. .664
BASE COUNT      174 a    161 c    191 g    152 t
ORIGIN

  Query Match           69.2%;  Score 376.2;  DB 5;  Length 678;
  Best Local Similarity   83.7%;  Pred. No. 2.3e-101;
  Matches  453;  Conservative    0;  Mismatches   78;  Indels   10;  Gaps    2;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              |||||||||||||||||| |||| || ||| |||||||||  |||||||||||||  ||
Db        146 GAAGGGCATTGCTACAAGGCCTTCGAAAAATACAAGACCTGGGAAGATGCAGAGAGGGTC 205

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||||| |||||| ||| || | ||||||||||||||||||||||||||||||||||||
Db        206 TGCACGGAGCAGGCGAAGGGCGCGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 265

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  ||||||| || |||  |  | |  || || ||||||||||
Db        266 TTTGTGGCCCAGCTGGTCACTCAGAACATGAAGAGACTCGATTTCTATATCTGGATCGGG 325

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| ||||  |||||  |||||||||| |  ||||||||||||||||||||||  ||
Db        326 CTGAGGGTTCAAGGCAAAGTCAAGCAATGCAACTCGGAGTGGAGCGATGGCTCCAGCGTC 385

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||| |||||||||| |||||||||| |||||||||  || | | |||||||
Db        386 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACGTGTCTTGGGCTGGAAAAAGAGACA 445

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              |  |||| |||||||||  ||| |||||||| ||||||| ||||||| ||||||||||||
Db        446 GATTTTCGTAAGTGGGTCAATATTTACTGTGGACAACAAAATCCTTTCGTCTGCGAGGCA 505

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              |||||||||||| ||||||  |||||||||||||||||| |||||||||          ||
Db        506 TAGTCTGAAGATGCAGCTGTCTGAAGTCTGGAGAAGCAAAGAAGCCCCC--------ACC 557

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              | |  ||   |  ||| |||||||||| ||||||||| ||||||||||||||||||||
Db        558 CAACACCCCACCTGCCCCAATCTCTGCTCTGCACCCTTCGCTCAACGGATGCTCTCTGTA 617

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGCCT 538
              |||||||||||| ||||| |||||||||||| |||   |||||||||||||||||||| |
Db        618 GCTGGATCTGGTTTTGCTACTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCCTAGCAT 677

Qy        539 G 539
              |
Db        678 G 678


RESULT 6
AB046491
LOCUS       AB046491                  683 bp    mRNA    linear   VRT 27-JAN-2001
DEFINITION  Trimeresurus flavoviridis mRNA for factor XI/factor X binding
            protein A chain, complete cds.
ACCESSION   AB046491
VERSION     AB046491.1  GI:12583676
KEYWORDS    .
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (sites)
  AUTHORS   Tani,A., Ogawa,T., Nose,T., Nikandorov,N.N., Deshimaru,M.,
            Chijiwa,T., Chang,C., Fukumaki,Y. and Ohno,M.
  TITLE     Characterization and molecular evolution of an anticoagulant
            protein from Agkistrodon actus venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 683)
  AUTHORS   Ogawa,T. and Ohno,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2000) Tomohisa Ogawa, Tohoku University, Dept. of
            Agricultural Science; 1-1 Tsutsumidori Amamiyamachi, Aoba-ku,
```

```
           Sendai, Miyagi 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
           Tel:81-22-717-8808, Fax:81-22-717-8807)
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
                     /tissue_type="venome gland"
                     /country="Japan:Kagoshima, Tokunoshima island"
     gene            1. .683
                     /gene="XI/Xbp-a"
     CDS             77. .517
                     /gene="XI/Xbp-a"
                     /codon_start=1
                     /product="factor XI/factor X binding protein A chain"
                     /protein_id="BAB21452.1"
                     /db_xref="GI:12583677"
                     /translation="MSFGLLVVAASLRGTGADCLSGWSSYEGHCYKAFEKYKTWEDAE
                     RVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIWIGLRVQGKVKQCNSEWSD
                     GSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCGQQNPFVCEA"
BASE COUNT       172 a    162 c    194 g    155 t
ORIGIN

  Query Match             65.2%;  Score 354.8;  DB 5;  Length 683;
  Best Local Similarity   82.7%;  Pred. No. 6e-95;
  Matches  445;  Conservative    0;  Mismatches   82;  Indels   11;  Gaps    3;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              || |||||||||||||||| |||| || |||  |||||||||  ||||||||||||  ||
Db        155 GAGGGGCATTGCTACAAGGCCTTCGAAAAATACAAGACCTGGGAAGATGCAGAGAGGGTC 214

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||||| |||||| ||| || | ||||||||||||||||||||||||||||||||||||
Db        215 TGCACGGAGCAGGCGAAGGGCGCGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 274

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  ||||||| || |||  |  | |  || || ||||||||||
Db        275 TTTGTGGCCCAGCTGGTCACTCAGAACATGAAGAGACTCGATTTCTATATCTGGATCGGG 334

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| ||||   || |  |||||||||| |  ||||||||||||||||||||| ||
Db        335 CTGAGGGTTCAAGGAAAGGTCAAGCAATGCAACTCGGAGTGGAGCGATGGCTCCAGCGTC 394

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||| |||||||||| |||||||||| |||||||||  || | | |||||||
Db        395 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACGTGTCTTGGGCTGGAAAAAGAGACA 454

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              |  |||| ||||||||  ||| |||||||| ||||||| ||||||| ||||||||||||
Db        455 GATTTTCGTAAGTGGGTCAATATTTACTGTGGACAACAAAATCCTTTCGTCTGCGAGGCA 514

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              |||||||||||| ||||||  |||||||||||||||||| |||||||||          ||
Db        515 TAGTCTGAAGATGCAGCTGTCTGAAGTCTGGAGAAGCAAAGAAGCCCCC--------ACC 566

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              | |  ||   |  ||| ||||||   || ||||||| ||||||||||||||||||||||
Db        567 CAACACCCCACCTGCCCCAATCTGCTCTCTGCACCC-TTGCTCAACGGATGCTCTCTGTA 625

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGC 536
              |||||||||||| ||||| |||||||||||||| ||||   ||||||||||||||||||
Db        626 GCTGGATCTGGTTTTGCTACTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCCTAGC 683


RESULT 7
AF540646
LOCUS       AF540646                661 bp    mRNA    linear   VRT 26-SEP-2002
DEFINITION  Deinagkistrodon acutus agglucetin-alpha 2 subunit precursor, mRNA,
            complete cds.
ACCESSION   AF540646
VERSION     AF540646.1  GI:23321260
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Wang,W.J. and Huang,T.F.
  TITLE     A novel tetrameric venom protein, agglucetin from Agkistrodon
            acutus, acts as a glycoprotein Ib agonist
  JOURNAL   Thromb. Haemost. 86 (4), 1077-1086 (2001)
  MEDLINE   21542539
   PUBMED   11686327
REFERENCE   2  (bases 1 to 661)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Molecular structure and functional characterization of agglucetin,
            a tetrameric glycoprotein Ib-binding protein, from Formosan pit
            viper
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 661)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
            of Technology, 261, Wen-Hwa 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
            333-03, R.O.C.
FEATURES             Location/Qualifiers
     source          1. .661
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
     CDS             13. .489
                     /note="platelet agglutination inducer; glycoprotein Ib
                     agonist"
                     /codon_start=1
                     /product="agglucetin-alpha 2 subunit precursor"
                     /protein_id="AAN23125.1"
                     /db_xref="GI:23321261"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADFNCPPGWSAYDQYCYQVIKE
                     PKNWDDAERFCTEQADGGHLVSIESKGERDFVAQLVSQNIESVEDHVWTGLRVQNKEK
                     QCSTEWSDGSSVSYENLLELYMRKCGALERETGFHKWINLGCIQLNPFVCKFPPQC"
     sig_peptide     13. .81
     mat_peptide     82. .486
                     /product="agglucetin-alpha 2 subunit"
BASE COUNT      167 a    159 c    183 g    152 t
ORIGIN

  Query Match             61.7%;  Score 335.8;  DB 5;  Length 661;
  Best Local Similarity   79.6%;  Pred. No. 2.9e-89;
  Matches  437;  Conservative    0;  Mismatches  102;  Indels   10;  Gaps    3;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||  | |||||||| ||||| ||||| || | || | ||||   |||||||| | |||
Db        115 GATCAGTATTGCTACCAGGTCATCAAAGAACCGAAAAACTGGGATGATGCAGAGCGGTTC 174

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||||| |||||| | |||| ||||| ||||||||||||||||||    |||||   ||||
Db        175 TGCACGGAGCAGGCGGACGGTGGGCACCTGGTCTCTATCGAAAGCAAGGGAGAGAGAGAC 234

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  ||||||| ||| ||    |   ||  |||||||||| ||||
Db        235 TTTGTGGCCCAGCTGGTCTCTCAGAACATAGAGAGTGTTGAAGACCATGTCTGGACCGGA 294

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| ||||||||||||||||||||||||| ||||||||||||||||||||| ||
Db        295 CTGAGGGTTCAAAACAAAGAAAAGCAATGCAGCACAGAGTGGAGCGATGGCTCCAGCGTC 354

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||   ||||| | |   | |||||||   | | || |   |||||||
Db        355 AGTTATGAGAACCTGCTTGAACTATATATGAGAAAGTGTGGGGCGCTGGAACGAGAGACA 414

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG--- 357
              ||||||||||||||   ||| ||  ||||  |||||   ||||||| |||||| ||
Db        415 GGGTTTCATAAGTGGATCAATCTTGGCTGTATACAACTGAATCCTTTCGTCTGCAAGTTC 474

Qy        358 ---GCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
```

```
                 || ||| | ||||||||||||| ||||||||||||||||||||||||||||||||
Db        475 CCACCACAGTGTTAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCA 534

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              ||   |||| |||   | ||| |||||||||||| ||||||||||||||| ||||||||
Db        535 CA--GCCCACCCCCTGCCTGCCGCAATCTCTGCTCTGCACCCTTTGCTCAATGGATGCTC 592

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||||||||||||| |||||||||||||||||||| ||||   |||||||||||||
Db        593 TCTGTAGCTGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCC 652

Qy        533 TAGCCTGAA 541
              ||||   ||
Db        653 TAGCAAAAA 661


RESULT 8
AF190827
LOCUS       AF190827                 630 bp    mRNA    linear   VRT 01-OCT-2000
DEFINITION  Gloydius halys halyxin A-chain precursor (HXNA) mRNA, complete cds.
ACCESSION   AF190827
VERSION     AF190827.1  GI:10441755
KEYWORDS    .
SOURCE      Gloydius halys (halys viper)
  ORGANISM  Gloydius halys
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Gloydius.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
  TITLE     A novel coagulation factor Xa inhibitor from Korean snake
            (Agkistrodon halys) venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 630)
  AUTHORS   Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-1999) Cardiovascular Center, Yonsei University
            College of Medicine, 134 Shinchon-dong, Seoul 120-752, Republic of
            Korea
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Gloydius halys"
                     /mol_type="mRNA"
                     /db_xref="taxon:8714"
     gene            1. .630
                     /gene="HXNA"
     CDS             24. .482
                     /gene="HXNA"
                     /function="coagulation factor Xa inhibitor"
                     /codon_start=1
                     /product="halyxin A-chain precursor"
                     /protein_id="AAG17178.1"
                     /db_xref="GI:10441756"
                     /translation="MGRFIFLSFGLLVVFLSLSGTGADCPSGWSSYEGHCYNIFHLFK
                     TWAEAERFCRKQVKGAHLVSIESSEEADFVAQLVSENMKRYGIYIWIGLRVRGKKKQC
                     SSQWSDGSSVSYQNWIEAESKTCLGLQKETEFRKWFNIYCGERNPFVCEA"
     sig_peptide     24. .92
                     /gene="HXNA"
     mat_peptide     93. .479
                     /gene="HXNA"
                     /product="halyxin A-chain"
BASE COUNT      173 a    137 c    174 g    146 t
ORIGIN

  Query Match             57.6%;  Score 313.4;  DB 5;  Length 630;
  Best Local Similarity   84.2%;  Pred. No. 1.5e-82;
  Matches  353;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              |||||||||||||||||  ||||| | | ||  |||||||| |||| ||||||| |||
Db        120 GAAGGGCATTGCTACAACATCTTCCATCTATTCAAGACCTGGGCAGAAGCAGAGAGGTTC 179

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
```

```
              |||| |||||||||||| || | |||||||||||||||||||||||||| ||||||||||
Db        180 TGCAGGAAGCAGGTGAAGGGCGCGCATCTGGTCTCTATCGAAAGCTCCGAAGAAGCAGAC 239

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  || |||| || |||  |  |  ||| || ||||||||||
Db        240 TTTGTGGCCCAGCTGGTCTCTGAGAACATGAAGAGATACGGGATCTATATCTGGATCGGA 299

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| || |  |||| | |||||||||||||     ||||||||||||||||||| ||
Db        300 CTGAGGGTTCGAGGCAAAAAGAAGCAATGCAGCTCCCAGTGGAGCGATGGCTCCAGCGTC 359

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||| |||| |||||||||| |||||||||| |||||||||  |||| | |||||||
Db        360 AGTTATCAGAACTGGATTGAAGCAGAATCCAAAACGTGTCTTGGGCTGCAAAAAGAGACA 419

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              | ||||| |||||||   ||| |||||||||  | ||| | |||||| ||||||||||
Db        420 GAGTTTCGTAAGTGGTTTAATATTTACTGTGGAGAACGAAATCCTTTCGTCTGCGAGGCA 479

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCC 419
              |||||||||||| ||||||  ||||||||||||||||||||||||||||||||||  ||
Db        480 TAGTCTGAAGATGCAGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCCACC 538


RESULT 9
AF540645
LOCUS       AF540645                 666 bp    mRNA    linear   VRT 26-SEP-2002
DEFINITION  Deinagkistrodon acutus agglucetin-alpha 1 subunit precursor, mRNA,
            complete cds.
ACCESSION   AF540645
VERSION     AF540645.1  GI:23321258
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 666)
  AUTHORS   Wang,W.J. and Huang,T.F.
  TITLE     A novel tetrameric venom protein, agglucetin from Agkistrodon
            acutus, acts as a glycoprotein Ib agonist
  JOURNAL   Thromb. Haemost. 86 (4), 1077-1086 (2001)
  MEDLINE   21542539
   PUBMED   11686327
REFERENCE   2  (bases 1 to 666)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Molecular structure and functional characterization of agglucetin,
            a tetrameric glycoprotein Ib-binding protein, from Formosan pit
            viper
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 666)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
            of Technology, 261, Wen-Hwa 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
            333-03, R.O.C.
FEATURES             Location/Qualifiers
     source          1. .666
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
     CDS             18. .482
                     /note="platelet agglutination inducer; glycoprotein Ib
                     agonist"
                     /codon_start=1
                     /product="agglucetin-alpha 1 subunit precursor"
                     /protein_id="AAN23124.1"
                     /db_xref="GI:23321259"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADVDCLPGWSAYDQSCYRVFKL
                     LKTWDDAEKFCTERPKGGHLVSIESAGERDFVAQLVSENKQTDNVWLGLKIQSKGQQC
                     STEWTDGSSVSYENFSEYQSKKCFVLEKNTGFRTWLNLNCGSEYAFVCKSPP"
     sig_peptide     18. .86
     mat_peptide     87. .479
```

/product="agglucetin-alpha 1 subunit"
BASE COUNT 166 a 163 c 178 g 159 t
ORIGIN

Query Match 57.4%; Score 312; DB 5; Length 666;
Best Local Similarity 78.7%; Pred. No. 4e-82;
Matches 428; Conservative 0; Mismatches 100; Indels 16; Gaps 4;

```
Qy          9 TTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGAA 68
              ||||||| |||||||||||     |||||||||  ||||||||| |||||||||  |
Db        128 TTGCTACCGGGTCTTCAAACTCCTCAAGACCTGGGATGATGCAGAGAAGTTCTGCACGGA 187

Qy         69 GCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGG 128
              || |  ||| || |||||||||||||||||||||||| | |||||   ||||||||||
Db        188 GCGGCCGAAAGGCGGGCATCTGGTCTCTATCGAAAGCGCGGGAGAGAGAGACTTTGTGGC 247

Qy        129 CCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGGC 188
              |||| || |  || |||| | | || | | ||       |||||||| | |||||| |
Db        248 CCAGCTGGTCTCTGAGAACAAACAGACGGACA------ATGTCTGGCTTGGACTGAAGAT 301

Qy        189 TCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGA 248
              ||||| ||||| | |||||||||||| |||||||| |||||||||||| ||||||||||
Db        302 TCAAAGCAAAGGACAGCAATGCAGCACAGAGTGGACCGATGGCTCCAGCGTCAGTTATGA 361

Qy        249 GAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAGGGTTTCA 308
              |||||    |||| |  |||||||||||||| ||| | || | | | | ||||||||||
Db        362 GAATTTTTCTGAATACCAATCCAAAAAGTGTTTTGTGCTGGAAAAAAACACAGGGTTTCG 421

Qy        309 TAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA------GGCATA 362
              | |||   ||| || | |||| |     || || |||| |||||| |       | ||||
Db        422 CACATGGCTCAATCTTAATTGTGGAAGTGAATATGCTTTCGTCTGCAAGTCCCCGCCATA 481

Qy        363 GTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCCCC 422
              || | |||||||||||||| |||||||||||||||||||||||||||||||||    |||
Db        482 GTGTTAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCA--GCCC 539

Qy        423 AACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTAGC 482
              | |||    |  ||| | ||||||||| ||||||||||||||||||||||||||||||
Db        540 ACCCCCTGCCTGCCGCTATCTCTGCTCTGCACCCTTTGCTCAACGGATGCTCTCTGTAGC 599

Qy        483 TGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGCCTGA 540
              |||||||||| |||||||||||||||||||| ||||   |||||||||||||||||    |
Db        600 TGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCCTAGCAAAA 659

Qy        541 AAAA 544
              ||||
Db        660 AAAA 663
```

RESULT 10
AX427204
LOCUS AX427204 721 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 4 from Patent WO0214364.
ACCESSION AX427204
VERSION AX427204.1 GI:21530566
KEYWORDS .
SOURCE Sistrurus miliarius (pigmy rattlesnake)
ORGANISM Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Sistrurus.
REFERENCE 1
AUTHORS Sheppard,P.O. and Bishop,P.D.
TITLE Rattlesnake venom gland proteins
JOURNAL Patent: WO 0214364-A 4 21-FEB-2002;
ZymoGenetics, Inc. (US)
FEATURES Location/Qualifiers
source 1. .721
/organism="Sistrurus miliarius"
/mol_type="genomic DNA"
/db_xref="taxon:8758"
CDS 91. .549
/note="Zsnk3"
/codon_start=1
/protein_id="CAD35499.1"
/db_xref="GI:21530567"
/translation="MGRFIFVSFGLLVVFLSLSGTGADCPSGWSSYDQHCYRVFKQLK
TWDDAERFCSEQAEGGHLVSIESSEEAAFVAQLVPENRRRAILYIWIGLRVQGKEKQC
SAKWSDGSSVSYENWIEAESKTCLGLQQGTNYHKWVNIYCGEINPFVCEA"
BASE COUNT 192 a 167 c 203 g 159 t
ORIGIN

Query Match 56.3%; Score 306.2; DB 6; Length 721;
Best Local Similarity 78.8%; Pred. No. 2.2e-80;
Matches 365; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   |||||||||| ||||||||||||   ||||| |||   |||||||||| |||
Db        187 GATCAGCATTGCTACAGGGTCTTCAAACAACTCAAGACGTGGGACGATGCAGAGAGGTTC 246

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| | | || |||||||| ||||||||||||||||||| |||||||| |
Db        247 TGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTCTCTATCGAAAGCTCCGAAGAAGCAGCC 306

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  || |||| |  | |  |||||   || || ||||||||||
Db        307 TTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGGAGAGCCATTCTCTATATCTGGATCGGA 366

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| ||||  |||||| ||||||||||||    ||||||||||||||||||| ||
Db        367 CTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGCGCGAAGTGGAGCGATGGCTCCAGCGTC 426

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||| |||||||||| |||||||||| |||||||| ||||   ||  |||
Db        427 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACATGTCTTGGGCTGCAACAAGGCACA 486

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              | |||||||||||  ||| |||||||| | ||   | ||||||||||||||||||||
Db        487 AATTATCATAAGTGGGTCAATATTTACTGTGGAGAAATAAATCCTTTTGTCTGCGAGGCA 546

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              |||||||||||| ||||||  |||||||| ||||||||||||||||||||||    ||
Db        547 TAGTCTGAAGATGCAGCTGTGTGAAGTCTGCAGAAGCAAGGAAGCCCCCCACCCCCCACC 606

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTC 463
              ||   |||||     |  | ||  | | ||    |   |||||
Db        607 CCCCACCTGCTGCATCTGTAGCTGGGATCTGGTTCTGCTGCTC 649
```

RESULT 11
AF102901
LOCUS AF102901 704 bp mRNA linear VRT 22-DEC-2000
DEFINITION Deinagkistrodon acutus agkicetin alpha subunit precursor (GPIbA1) mRNA, complete cds.
ACCESSION AF102901
VERSION AF102901.1 GI:11967284
KEYWORDS .
SOURCE Deinagkistrodon acutus
ORGANISM Deinagkistrodon acutus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE 1 (bases 1 to 704)
AUTHORS Chen,Y.L., Hong,T.M., Chang,T. and Tsai,I.H.
TITLE cDNA sequence and functional characterization of glycoprotein Ib-binding protein from the venom of Deinagkistrodon acutus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 704)
AUTHORS Chen,Y.L. and Tsai,I.H.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-1998) Biological Chemistry, Academia Sinica, Taipei, Taiwan, R.O.C.
FEATURES Location/Qualifiers
source 1. .704
/organism="Deinagkistrodon acutus"
/mol_type="mRNA"

```
                     /db_xref="taxon:36307"
     gene            <1. .704
                     /gene="GPIbA1"
     CDS             61. .528
                     /gene="GPIbA1"
                     /note="antagonist of vWF-induced platelet agglutination"
                     /codon_start=1
                     /product="agkicetin alpha subunit precursor"
                     /protein_id="AAG42040.1"
                     /db_xref="GI:11967285"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCLPGWSSYIRFCYQPFKLLK
                     TWEDAERFCTEQANGGHLVSFESAREADFVAGVLSENIKIKPYVWIGLRVQNEGQQCS
                     SKWSDSSKVSYENLVEPFSKKCFVLKKDTGFRTWENVYCGLKHVFMCKYLKPR"
     sig_peptide     61. .129
                     /gene="GPIbA1"
     mat_peptide     130. .525
                     /gene="GPIbA1"
                     /product="agkicetin alpha subunit"
BASE COUNT      175 a    167 c    187 g    175 t
ORIGIN

  Query Match             56.2%;  Score 306;  DB 5;  Length 704;
  Best Local Similarity   77.7%;  Pred. No. 2.5e-80;
  Matches  426;  Conservative    0;  Mismatches  105;  Indels   17;  Gaps    4;

Qy          5 GGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCA 64
              ||  ||||||| ||  |||||| |      |||||||| ||||||||||||| |||||||
Db        161 GGTTTTGCTACCAGCCCTTCAAGCTCCTCAAGACCTGGGAAGATGCAGAGAGGTTCTGCA 220

Qy         65 CGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTG 124
              || |||||| ||| |||||||||||||||||| |||||||| || ||||||||||||||
Db        221 CGGAGCAGGCGAATGGGGGGCATCTGGTCTCTTTCGAAAGCGCCAGAGAAGCAGACTTTG 280

Qy        125 TGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGA 184
              ||| |  | || |  || |||| |||||    |  ||||  | ||||||||||  |||||||
Db        281 TGGCCGGGGTGCTCTCTGAGAACATAAA---AATCAAACCCTATGTCTGGATTGGACTGA 337

Qy        185 GGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTT 244
              ||| |||||||  ||| | ||||||||||||     |||||||  ||| ||||||    ||||||
Db        338 GGGTTCAAAACGAAGGACAGCAATGCAGCTCTAAGTGGAGTGATAGCTCCAAAGTCAGTT 397

Qy        245 ATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAGGGT 304
              |||||||  | | |||||   |   |||||||||||| ||| | || | |  || |||||||
Db        398 ATGAGAACTTGGTTGAACCATTTTCCAAAAAGTGTTTTGTGCTGAAAAAGGACACAGGGT 457

Qy        305 TTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA------GG 358
              ||| || |||||||||| |||||||| || | || ||  |||  | ||| |       |
Db        458 TTCGTACGTGGGAGAATGTTTACTGTGGACTAAAACATGTTTTCATGTGCAAGTACCTGA 517

Qy        359 CATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATC 418
              |  |  | ||||||  ||||   |||||||||||||||||||||||||||||||||  |
Db        518 AACCGCGTTAAGATCTGGCTGTGCGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCCAC 577

Qy        419 CCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTG 478
              ||||  ||       ||| | || ||| || |||||||||||||||||||||||||||||
Db        578 CCCCCGCC------TGCCGCTATGTCTTCTCTGCACCCTTTGCTCAACGGATGCTCTCTG 631

Qy        479 TAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGC 536
              |||||||||||||| |||||||||||||||||||| ||||  ||||||||||||||||||
Db        632 TAGCTGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCCTAGC 691

Qy        537 CTGAAAAA 544
                  |||||
Db        692 AAAAAAAA 699


RESULT 12
AY149341
LOCUS       AY149341                681 bp    mRNA    linear   VRT 24-NOV-2002
DEFINITION  Trimeresurus flavoviridis flavocetin-A alpha chain mRNA, complete
            cds.
ACCESSION   AY149341
VERSION     AY149341.1  GI:25245560
KEYWORDS    .
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   Shin,Y., Okuyama,I., Hasegawa,J. and Morita,T.
  TITLE     Molecular cloning of glycoprotein Ib-binding protein, flavocetin-A,
            which inhibits platelet aggregation
  JOURNAL   Thromb. Res. 99 (3), 239-247 (2000)
  MEDLINE   20402624
   PUBMED   10942790
REFERENCE   2  (bases 1 to 681)
  AUTHORS   Shin,Y., Okuyama,I., Hasegawa,J. and Morita,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-SEP-2002) Department of Biochemistry, Meiji
            Pharmaceutical University, 2-522-1, Noshio, Kiyose, Tokyo 204-8588,
            Japan
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
                     /tissue_type="venom glands"
     CDS             37. .513
                     /note="partial peptide sequencing"
                     /codon_start=1
                     /evidence=experimental
                     /product="flavocetin-A alpha chain"
                     /protein_id="AAN72438.1"
                     /db_xref="GI:25245561"
                     /translation="MERLIFVSFGLLVVILSLSGTGADFDCIPGWSAYDRYCYQAFSK
                     PKNWEDAESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNKEQ
                     QCRSEWSDASSVNYENLVKQFSKKCYALKKGTELRTWFNVYCGTENPFVCKYTPEC"
BASE COUNT      170 a    164 c    182 g    165 t
ORIGIN

  Query Match             54.7%;  Score 297.4;  DB 5;  Length 681;
  Best Local Similarity   76.7%;  Pred. No. 9.4e-78;
  Matches  418;  Conservative    0;  Mismatches  116;  Indels   11;  Gaps    4;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||  || |||||||| ||| ||||||   || | || | ||||  |||||||||||| |||
Db        139 GATCGGTATTGCTACCAGGCCTTCAGTAAACCGAAAAACTGGGAAGATGCAGAGAGTTTC 198

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||  | ||  |||||||       |||| ||||||| ||||||||||||||||||| ||||
Db        199 TGCGAGGAGGGGGTGAAGACCTCGCATTTGGTCTCAATCGAAAGCTCCGGAGAAGGAGAC 258

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              || |||| |||| || | ||| |||||||||| | ||        |||||||||| |||
Db        259 TTCGTGGCCCAGCTGGTCGCTGAGAAGATAAAGACTTCCTTTCAGTATGTCTGGATTGGA 318

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||  |||||||||||| ||||||||||   |||||||||||| |||||||  ||
Db        319 CTGAGGATTCAAAACAAAGAACAGCAATGCAGGTCGGAGTGGAGCGATGCCTCCAGTGTC 378

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              | ||||||||| | | || || ||    |||||||| |||  || | || | | || ||||
Db        379 AATTATGAGAACTTGGTTAAACAATTTTCCAAAAAATGTTATGCGCTGAAAAAAGGGACA 438

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              | | |||  | ||||   ||  | ||||||| |  | || ||||||| |||||| |
Db        439 GAGCTTCGCACGTGGTTCAACGTATACTGTGGAACAGAAAATCCTTTCGTCTGCAAGTAC 498

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                | || | | | |||||||||||| |||||||||||||||||||||||||||||| |||
Db        499 ACGCCAGAATGTTAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCGCCC 558

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473
              ||  ||||| |||   | ||| |||||||||| ||| |||| |   || | |||||||
Db        559 CA--CCCCACCCCCTGCCTGCCGCAATCTCTGCTCTGCCCCCTGTCATTCCATGGATGCT 616
```

```
Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
              ||||||||||||||||||| |||||||||||||||||| ||||   ||||||||||||
Db        617 CTCTGTAGCTGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 676

Qy        532 CTAGC 536
              |||||
Db        677 CTAGC 681
```

RESULT 13
CDRNACVXA
LOCUS CDRNACVXA 800 bp mRNA linear VRT 30-JUL-1998
DEFINITION Crotalus durissus mRNA for convulxin alpha.
ACCESSION Y16348
VERSION Y16348.1 GI:3367645
KEYWORDS convulxin alpha; CVX alpha gene.
SOURCE Crotalus durissus (tropical rattlesnake)
ORGANISM Crotalus durissus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Crotalus.
REFERENCE 1
AUTHORS Leduc,M. and Bon,C.
TITLE Cloning of subunits of convulxin, a collagen-like platelet-aggregating protein from Crotalus durissus terrificus venom
JOURNAL Biochem. J. 333 (Pt 2), 389-393 (1998)
MEDLINE 98324901
PUBMED 9657980
REFERENCE 2 (bases 1 to 800)
AUTHORS Leduc,M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1998) M. Leduc, Institut Pasteur, Unite des Venins, 25 Rue du Dr Roux, 75724 Paris cedex 15, Paris, FRANCE

FEATURES	Location/Qualifiers
source	1. .800
	/organism="Crotalus durissus"
	/mol_type="mRNA"
	/sub_species="terrificus"
	/db_xref="taxon:8731"
	/clone="pT3T7"
	/tissue_type="venom gland"
gene	1. .800
	/gene="CVX alpha"
mRNA	1. .800
	/gene="CVX alpha"
	/product="convulxin alpha"
CDS	133. .609
	/gene="CVX alpha"
	/codon_start=1
	/evidence=experimental
	/product="convulxin alpha"
	/protein_id="CAA76181.1"
	/db_xref="GI:3367646"
	/db_xref="SPTREMBL:O93426"
	/translation="MGRFIFVSFGLLVLFLSLSGTGAGLHCPSDWYYYDQHCYRIFNE EMNWEDAEWFCTKQAKGAHLVSIKSAKEADFVAWMVTQNIEESFSHVSIGLRVQNKEK QCSTKWSDGSSVSYDNLLDLYITKCSLLKKETGFRKWFVASCIGKIPFVCKFPPQC"
mat_peptide	211. .606
	/gene="CVX alpha"
	/product="convulxin alpha"
	/evidence=experimental

BASE COUNT 203 a 183 c 218 g 196 t
ORIGIN

Query Match 53.0%; Score 288.2; DB 5; Length 800;
Best Local Similarity 75.2%; Pred. No. 5.4e-75;
Matches 416; Conservative 0; Mismatches 123; Indels 14; Gaps 4;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   ||||||||||  | |||||||  ||    | || ||||  |||||||||| | |||
Db        235 GATCAGCATTGCTACCGGATCTTCAATGAAGAGATGAACTGGGAAGATGCAGAGTGGTTC 294

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||||||||||| ||| || | |||||| |||||||| ||||| ||  ||||||||||
Db        295 TGCACGAAGCAGGCGAAGGGCGCGCATCTCGTCTCTATCAAAAGCGCCAAAGAAGCAGAC 354

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |  | ||| |  ||||||| ||| ||  | ||      |||||||| ||| |||
Db        355 TTTGTGGCCTGGATGGTCACTCAGAACATAGAGGAATCCTTTTCCCATGTCTCGATTGGA 414

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| |||||||||||||||||||||||||||    ||||||||||||||||||| ||
Db        415 CTGAGGGTTCAAAACAAAGAAAAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 474

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||| || | | | ||  | |    | |||||||  |  | || | | |||||||
Db        475 AGTTATGACAACTTGTTGGATCTATATATTACAAAGTGTAGTCTGCTGAAAAAAGAGACA 534

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              ||||||| |||||||      |  | |||| |  | |  ||||| ||||||  |
Db        535 GGGTTTCGTAAGTGGTTCGTTGCTAGCTGTATAGGAAAGATTCCTTTCGTCTGCAAGTTC 594

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                | || ||| | |||  ||||||   |||||||||||||||||||  |||||| |
Db        595 CCGCCACAGTGTTAAGTACCAGCTTTGTGAAGTCTGGAGAAGCAAGGAGCCCCCCCCC-- 652

Qy        415 CATCCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473
                  ||||| ||| |||  ||| |||||||||| |||||||| |   || | || ||||
Db        653 ---TCCCCACCCCCGCCCTGCCGCAATCTCTGCTCTGCACCCTGTCATTCCATGGGTGCT 709

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
              |||||| ||||||||||||  |||||||||||||||||| ||||   ||||||||||||
Db        710 CTCTGTGGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 769

Qy        532 CTAGCCTGAAAAA 544
              ||||| |||||||
Db        770 CTAGCATGAAAAA 782
```

RESULT 14
AX427210
LOCUS AX427210 725 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 10 from Patent WO0214364.
ACCESSION AX427210
VERSION AX427210.1 GI:21530572
KEYWORDS .
SOURCE Sistrurus miliarius (pigmy rattlesnake)
ORGANISM Sistrurus miliarius
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Sistrurus.
REFERENCE 1
AUTHORS Sheppard,P.O. and Bishop,P.D.
TITLE Rattlesnake venom gland proteins
JOURNAL Patent: WO 0214364-A 10 21-FEB-2002; ZymoGenetics, Inc. (US)

FEATURES	Location/Qualifiers
source	1. .725
	/organism="Sistrurus miliarius"
	/mol_type="genomic DNA"
	/db_xref="taxon:8758"
CDS	88. .564
	/note="Zsnk5"
	/codon_start=1
	/protein_id="CAD35501.1"
	/db_xref="GI:21530573"
	/translation="MGRFIFVSFGLLVVFLSLSGTGADFNCPSGWFAYDQYCYRVIKR LKTWDDAERFCSEQAKGGHLASVENDEEAVFLAQLVAANIKQNQYYVWIGLRIQNKGQ QCSTKWSDGSSVSYENLVKSHSKKCFGLKKETEFLQWYNTDCEEKNLFVCKFPPEC"

BASE COUNT 179 a 171 c 200 g 175 t
ORIGIN

Query Match 52.9%; Score 288; DB 6; Length 725;
Best Local Similarity 75.9%; Pred. No. 6.2e-75;

Matches 416; Conservative 0; Mismatches 110; Indels 22; Gaps 4;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
Db        190 GATCAGTATTGCTACAGGGTCATCAAACGACTCAAGACCTGGGACGATGCAGAGCGGTTC 249

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
Db        250 TGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCGTCTGTCGAAAACGATGAAGAAGCAGTC 309

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
Db        310 TTTCTGGCCCAGTTGGTCGCTGCGAACATAAAGCAAAACCAATACTATGTCTGGATTGGA 369

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
Db        370 CTGAGGATTCAAAACAAAGGACAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 429

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
Db        430 AGTTATGAGAACCTGGTTAAATCACATTCCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 489

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
Db        490 GAGTTTCTTCAATGGTACAATACTGACTGCGAAGAAAAAAACCTTTTCGTCTGCAAGTTC 549

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
Db        550 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAATCCCCCCCCAC 609

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473
Db        610 CCGCC-------------TGCCACAATCTCTGCTCTGCACTCTGTCATTCCATGGATGCT 656

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
Db        657 CTCTGTGGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 716

Qy        532 CTAGCCTG 539
Db        717 CTAGCATG 724
```

```
RESULT 15
AX427207
LOCUS       AX427207                580 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 7 from Patent WO0214364.
ACCESSION   AX427207
VERSION     AX427207.1  GI:21530569
KEYWORDS    .
SOURCE      Sistrurus miliarius (pigmy rattlesnake)
  ORGANISM  Sistrurus miliarius
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Sistrurus.
REFERENCE   1
  AUTHORS   Sheppard,P.O. and Bishop,P.D.
  TITLE     Rattlesnake venom gland proteins
  JOURNAL   Patent: WO 0214364-A 7 21-FEB-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Sistrurus miliarius"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8758"
     CDS             <3. .437
                     /note="Zsnk4"
                     /codon_start=1
                     /protein_id="CAD35500.1"
                     /db_xref="GI:21530570"
                     /translation="IRNEGGTGADFDCPSDWYAYDQYCYRVIKQLRTWDDAERFCSEQ
                     AKGGHLVSIESDGEAAFVAQLVAENIKQNKYDVWIGLRIQGEEKQCSTKWSDGSSVNY
                     ENLIKHATKKCFGLKKETGFRTWRNVHCTQQNLFMCKFPPEC"
```

```
BASE COUNT      154 a    133 c    163 g    130 t
ORIGIN
```

Query Match 52.5%; Score 285.8; DB 6; Length 580;
Best Local Similarity 75.2%; Pred. No. 2.8e-74;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
Db         63 GATCAGTATTGCTACAGGGTCATCAAACAACTCAGGACGTGGGACGATGCAGAGAGGTTC 122

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
Db        123 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTCTCTATTGAAAGCGACGGAGAAGCAGCC 182

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
Db        183 TTTGTGGCCCAGCTGGTCGCTGAGAACATCAAGCAAAACAAATATGATGTCTGGATCGGA 242

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
Db        243 CTGAGGATTCAAGGCGAAGAGAAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 302

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
Db        303 AATTATGAGAACCTGATTAAACATGCGACCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 362

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
Db        363 GGGTTTCGCACGTGGCGCAATGTTCACTGTACACAACAAAATCTTTTCATGTGCAAGTTC 422

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
Db        423 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 482

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
Db        483 CACCGC------------------------------CACCCTTTGCTCAACGGATGCTC 511

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
Db        512 TCTGTAGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAACTCTTCC 571

Qy        533 TAGCCTGAA 541
Db        572 TAGCCTGGA 580
```

Search completed: December 8, 2003, 12:01:59
Job time : 3183 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 09:08:47 ; Search time 321 Seconds
(without alignments)
4574.749 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTACAAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	544	100.0	544	25	ABX93668	cDNA encoding Dein
2	306.2	56.3	721	24	AAD32055	Pigmy rattlesnake
3	288	52.9	725	24	AAD32059	Pigmy rattlesnake
4	285.8	52.5	580	24	AAD32057	Pigmy rattlesnake
5	276.2	50.8	632	24	AAL42015	Korean adder snake
6	266	48.9	601	22	AAI71876	Snake venom blood
7	226.4	41.6	690	16	AAQ89309	Snake venom antith
8	226.4	41.6	690	21	AAC61144	DNA encoding a sna
9	226.4	41.6	690	24	AAK99834	DNA encoding the a
10	186.6	34.3	456	24	AAD32056	Pigmy rattlesnake
11	175	32.2	432	24	AAD32058	Pigmy rattlesnake
12	166	30.5	474	24	AAD32060	Pigmy rattlesnake
13	144	26.5	633	22	AAI71877	Snake venom blood
14	102.6	18.9	272	16	AAQ89308	Snake venom antith
15	102.6	18.9	272	21	AAC61143	Rattlesnake venom
16	98.8	18.2	583	24	AAL42016	Korean adder snake
17	82.4	15.1	483	24	AAD32053	Pigmy rattlesnake
18	81.2	14.9	716	17	AAT64829	Snake venom protei
19	64.2	11.8	453	24	AAD32054	Pigmy rattlesnake
20	58	10.7	454	25	ABX93674	cDNA encoding Dein
c 21	39	7.2	10598	23	ABL27241	Drosophila melanog
c 22	39	7.2	10629	22	AAH28279	Nucleotide sequenc
23	39	7.2	15667	23	ABL27240	Drosophila melanog
24	36.8	6.8	1597	21	AAC59784	Human secreted pro
25	36.8	6.8	1597	22	AAF24170	Human secreted pro
c 26	36.8	6.8	128600	24	ABK83461	Human cDNA differe
27	36.4	6.7	498	24	ABK46204	cDNA encoding colo
28	36	6.6	492	22	AAI64154	Human Zlec1 encodi
29	36	6.6	492	24	AAD27975	Human lectin Zlec1
30	34.8	6.4	461	24	ABK46210	cDNA encoding colo
c 31	34.8	6.4	509	24	ABV97585	Human pancreatic c
32	34.6	6.4	2949	21	AAD00024	Degenerate DNA enc
33	34.6	6.4	2949	24	AAD37455	Human Zfsta2 degen
34	34.4	6.3	1221	23	AAS67247	DNA encoding novel
35	34.4	6.3	1222	23	AAS66042	DNA encoding novel
36	34.4	6.3	2581	23	AAS64652	DNA encoding novel
37	34.2	6.3	441	11	AAQ05622	Sequence encoding
38	34.2	6.3	493	24	ABK46235	cDNA encoding colo
39	34.2	6.3	495	10	AAN91109	Human reg cDNA. H
40	34.2	6.3	498	9	AAN81962	Sequence of human
41	34.2	6.3	800	22	AAF75077	Human colon associ
42	34.2	6.3	843	21	AAC77946	Human cancer assoc
43	34.2	6.3	843	22	AAF75090	Human colon associ
44	34	6.2	790	11	AAQ05300	Bovine pancreatic
45	33.8	6.2	585	24	ABK46227	cDNA encoding colo

ALIGNMENTS

RESULT 1
ABX93668
ID ABX93668 standard; cDNA; 544 BP.
XX
AC ABX93668;
XX
DT 02-JUN-2003 (first entry)
XX
DE cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
KW Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
KW platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW thrombosis.
XX
OS Deinagkistrodon acutus.
XX
FH Key Location/Qualifiers
FT CDS 1..363
FT /*tag= a
FT /product= "Antithrombosis enzyme alpha chain"
FT /partial
FT /transl_except= (pos:244..246,aa:Lys)
FT /note= "No start codon given"
XX
PN US6489451-B1.
XX

PD 03-DEC-2002.
XX
PF 10-APR-1998; 98US-0058740.
XX
PR 10-APR-1997; 97US-043886P.
XX
PA (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI Li BX, Cheng X;
XX
DR WPI; 2003-352116/33.
DR P-PSDB; ABU08798.
XX
PT New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT preventing and treating vaso-occulusive and thromboembolic disorders,
PT including myocardial infarction, restenosis, cerebral thrombosis and
PT unstable angina -
XX
PS Example 3; Column 25-26; 19pp; English.
XX
CC The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC prevents platelet aggregation. The anti-thrombosis enzyme was
CC administered to rabbits intravenously. Thrombosis was determined before
CC and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC following administration at 0.005 micro/kg and this activity was
CC increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC composition is useful for preventing and treating vaso-occulusive and
CC thromboembolic disorders, including myocardial infarction, restenosis,
CC angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC peripheral arterial occulsion, stroke. It is also useful for treating
CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC diseases. The present sequence represents cDNA encoding the
CC Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
SQ Sequence 544 BP; 158 A; 122 C; 143 G; 121 T; 0 other;

Query Match 100.0%; Score 544; DB 25; Length 544;
Best Local Similarity 100.0%; Pred. No. 4.4e-161;
Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540

Qy        541 AAAA 544
              ||||
Db        541 AAAA 544
```

RESULT 2
AAD32055
ID AAD32055 standard; DNA; 721 BP.
XX
AC AAD32055;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk3 gene.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; gene; Zsnk3; ds.
XX
OS Sistrurus miliarius.
XX
FH Key Location/Qualifiers
FT CDS 91..549
FT /*tag= a
FT /product= "Zsnk3 protein"
FT sig_peptide 91..159
FT /*tag= b
FT mat_peptide 160..546
FT /*tag= c
FT /product= "Mature Zsnk3 protein"
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR P-PSDB; AAE20179.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Claim 5; Page 73-74; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the

CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk3 gene.
XX
SQ Sequence 721 BP; 192 A; 167 C; 203 G; 159 T; 0 other;

Query Match 56.3%; Score 306.2; DB 24; Length 721;
Best Local Similarity 78.8%; Pred. No. 5e-86;
Matches 365; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   |||||||||||| ||||||||||||    |||||  |||   ||||||||||| |||
Db        187 GATCAGCATTGCTACAGGGTCTTCAAACAACTCAAGACGTGGGACGATGCAGAGAGGTTC 246

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || ||||||  | | || ||||||||| ||||||||||||||||||| ||||||||  |
Db        247 TGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTCTCTATCGAAAGCTCCGAAGAAGCAGCC 306

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  ||||  || |  ||  |||| |  | |  |||||   || ||  ||||||||||
Db        307 TTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGGAGAGCCATTCTCTATATCTGGATCGGA 366

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||||  ||||  |||||  |||||||||||     ||||||||||||||||||||  ||
Db        367 CTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGCGCGAAGTGGAGCGATGGCTCCAGCGTC 426

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||||  |||||||||  ||||||||||   |||||||| ||||    ||  |||
Db        427 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACATGTCTTGGGCTGCAACAAGGCACA 486

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
                | ||||||||||  ||| |||||||||  |  ||  | ||||||||||||||||||||
Db        487 AATTATCATAAGTGGGTCAATATTTACTGTGGAGAAATAAATCCTTTTGTCTGCGAGGCA 546

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              ||||||||||| ||||||  |||||||||  ||||||||||||||||||||||||    ||
Db        547 TAGTCTGAAGATGCAGCTGTGTGAAGTCTGCAGAAGCAAGGAAGCCCCCCACCCCCCACC 606

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTC 463
              ||   |||||     |  | ||  | | ||   |  |||||
Db        607 CCCCACCTGCTGCATCTGTAGCTGGGATCTGGTTCTGCTGCTC 649
```

RESULT 3
AAD32059
ID AAD32059 standard; DNA; 725 BP.
XX
AC AAD32059;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk5 gene.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; gene; Zsnk5; ds.
XX
OS Sistrurus miliarius.
XX

FH	Key	Location/Qualifiers
FT	CDS	88..564
FT		/*tag= a
FT		/product= "Zsnk5 protein"
FT	sig_peptide	88..156
FT		/*tag= b
FT	mat_peptide	157..561
FT		/*tag= c
FT		/product= "Mature Zsnk5 protein"

XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR P-PSDB; AAE20181.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Claim 5; Page 77; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk5 gene.
XX
SQ Sequence 725 BP; 179 A; 171 C; 200 G; 175 T; 0 other;

Query Match 52.9%; Score 288; DB 24; Length 725;
Best Local Similarity 75.9%; Pred. No. 2.8e-80;
Matches 416; Conservative 0; Mismatches 110; Indels 22; Gaps 4;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   | |||||||||  |||| ||||||  |   |||||||||   ||||||||| | |||
Db        190 GATCAGTATTGCTACAGGGTCATCAAACGACTCAAGACCTGGGACGATGCAGAGCGGTTC 249

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || ||||||  ||| || ||||||||||   ||| ||||||  |    | ||||||||  |
Db        250 TGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCGTCTGTCGAAAACGATGAAGAAGCAGTC 309

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||| ||| |||||||  | |||   ||| ||||||  |   | ||  | ||||||||||  |||
Db        310 TTTCTGGCCCAGTTGGTCGCTGCGAACATAAAGCAAAACCAATACTATGTCTGGATTGGA 369

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||  |||||||||||  | ||||||||||||   ||||||||||||||||||||  ||
Db        370 CTGAGGATTCAAAACAAAGGACAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 429

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||    | || ||  | | ||||||||||| |||||  || | | |||||||
Db        430 AGTTATGAGAACCTGGTTAAATCACATTCCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 489

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              | ||||| | | ||| | |||  | |||| ||| || || | | ||| ||||||  |
Db        490 GAGTTTCTTCAATGGTACAATACTGACTGCGAAGAAAAAAACCTTTTCGTCTGCAAGTTC 549

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                | || ||| | ||||||| ||||  |||||||||||||||||||||| |||||| | |
```

```
Db        550 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAATCCCCCCCCCAC 609

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473
              |  ||                |||||||||||||||| ||||| || |   || | |||||||
Db        610 CCGCC-------------TGCCACAATCTCTGCTCTGCACTCTGTCATTCCATGGATGCT 656

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
              |||||| ||||||||||| |||||||||||||||||| ||||   |||||||||||||
Db        657 CTCTGTGGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 716

Qy        532 CTAGCCTG 539
              |||||  ||
Db        717 CTAGCATG 724
```

RESULT 4
AAD32057
ID AAD32057 standard; DNA; 580 BP.
XX
AC AAD32057;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk4 gene.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; gene; Zsnk4; ds.
XX
OS Sistrurus miliarius.
XX
FH Key   Location/Qualifiers
FT CDS   3..437
FT   /*tag= a
FT   /product= "Zsnk4 protein"
FT   /note= "CDS does not include start codon"
FT   /partial
FT sig_peptide   3..29
FT   /*tag= b
FT mat_peptide   30..434
FT   /*tag= c
FT   /product= "Mature Zsnk4 protein"
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR P-PSDB; AAE20180.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Claim 5; Page 75; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk4 gene.
XX
SQ Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;

```
  Query Match             52.5%;  Score 285.8;  DB 24;  Length 580;
  Best Local Similarity   75.2%;  Pred. No. 1.2e-79;
  Matches  413;  Conservative    0;  Mismatches   97;  Indels   39;  Gaps    3;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   | ||||||||| |||| ||||||||   | ||| |||   |||||||||| |||
Db         63 GATCAGTATTGCTACAGGGTCATCAAACAACTCAGGACGTGGGACGATGCAGAGAGGTTC 122

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| ||| || |||||||| |||||||| ||||||  |||||||||| |
Db        123 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTCTCTATTGAAAGCGACGGAGAAGCAGCC 182

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || | ||| |||| || |||  | ||||   ||||||||||||||
Db        183 TTTGTGGCCCAGCTGGTCGCTGAGAACATCAAGCAAAACAAATATGATGTCTGGATCGGA 242

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||  ||||  | |||| |||||||||||||   ||||||||||||||||||||| ||
Db        243 CTGAGGATTCAAGGCGAAGAGAAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 302

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              | |||||||||   |||| || | |  ||||||||||| ||||| || | | |||||||
Db        303 AATTATGAGAACCTGATTAAACATGCGACCAAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 362

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              |||||||  | ||||  ||| || |||||  ||||||| ||| |||  | ||| |
Db        363 GGGTTTCGCACGTGGCGCAATGTTCACTGTACACAACAAAATCTTTTCATGTGCAAGTTC 422

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                | || ||| | ||||||| ||||  ||||||||||||||||||||||||||||||||
Db        423 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 482

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              || | |                                 ||||||||||||||||||||||
Db        483 CACCGC------------------------------CACCCTTTGCTCAACGGATGCTC 511

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||||||||||||  |||||||||||||||||| ||||   ||||||| ||| ||
Db        512 TCTGTAGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAACTCTTCC 571

Qy        533 TAGCCTGAA 541
              ||||||| |
Db        572 TAGCCTGGA 580
```

RESULT 5
AAL42015
ID AAL42015 standard; cDNA; 632 BP.
XX
AC AAL42015;
XX
DT 16-MAY-2002 (first entry)
XX
DE Korean adder snake venom salmorin A chain protein cDNA sequence.
XX
KW Korean adder; ss; salmorin protein A chain; snake venom;
KW fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW thrombin binding; blood coagulation.
XX

```
OS   Agkistrodon halys brevicaudus.
XX
FH   Key             Location/Qualifiers
FT   CDS             34..498
FT                   /*tag=  a
FT                   /product= "Salmorin A chain protein"
FT   sig_peptide     34..102
FT                   /*tag=  b
FT                   /note= "Signal peptide"
FT   mat_peptide     103..495
FT                   /*tag=  c
FT                   /note= "Mature Salmorin A chain protein"
FT   3'UTR           499..622
FT                   /*tag=  d
FT   polyA_signal    623..632
FT                   /*tag=  e
XX
PN   WO200214514-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-KR01277.
XX
PR   27-JUL-2000; 2000KR-0043470.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Chung K,  Kim D,  Koh Y;
XX
DR   WPI; 2002-241907/29.
DR   P-PSDB; AAO14520.
XX
PT   New salmorin protein derived from venom of Korean adder Agkistrodon
PT   halys brevicaudus, useful for treating thrombosis by repressing
PT   fibrinogen clotting through repression of activation of prothrombin
PT   into thrombin  -
XX
PS   Claim 1; Fig 1A; 30pp; English.
XX
CC   The invention comprises the nucleotide and protein sequences of a
CC   salmorin protein derived from the venom of Korean adder. Salmorin protein
CC   is composed of an A chain and a B chain, and has inhibitory activity
CC   against fibrinogen clotting. The salmorin protein of the invention is
CC   useful for treating thrombosis, as it represses fibrinogen clotting
CC   potently by binding to prothrombin and thrombin so as to delay blood
CC   coagulation. The present cDNA sequence encodes the Korean adder salmorin
CC   A chain protein.
XX
SQ   Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;

  Query Match             50.8%;  Score 276.2;  DB 24;  Length 632;
  Best Local Similarity   80.0%;  Pred. No. 1.4e-76;
  Matches  337;  Conservative    0;  Mismatches   83;  Indels    1;  Gaps    1;

Qy          2 AAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCT 61
              | ||  |||||||||  |||  ||||||  |||    | |||||||  ||||||||||| ||||
Db        137 ATGGTCATTGCTACCAGGCCTTCAATCAACGGATGACCTGGGAAGATGCAGAGAGGTTCT 196

Qy         62 GCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACT 121
              ||  |    |||| ||| || ||||||||||||||||||||||      | ||||||||||
Db        197 GCTCAGCACAGGCGAAGGGCGGGCATCTGGTCTCTATCGAAACACGGGCAGAAGCAGACT 256

Qy        122 TTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGAC 181
              ||||||  |||   ||  | |||  ||| |||| ||  ||  ||      ||||||||||||||||||
Db        257 TTGTGGCCCACGTGGTCGCTGAGAGGATAGAGACATCCTTTCCCCATGTCTGGATCGGAC 316

Qy        182 TGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCA 241
              ||||||  |  ||  |||||||  ||||||||||   ||||||||||||||  ||||||  |||
Db        317 TGAGGGATGAAGGCAAAGAACAGCAATGCAGCTCTGAGTGGAGCGATGGATCCAGCGTCA 376

Qy        242 GTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAG 301
              ||||||||||  ||||||||||  ||||||||||   |||||||||  ||  |   ||||   ||
Db        377 GTTATGAGAACTGGATTGAAGCAGAATCCAAAACATGTCTTGGGCTGGAACTAGACTCAA 436

Qy        302 GGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCAT 361
              |  ||||||||||||  |||  ||||||||||  ||||   |  |||||||  |||||||||||  |
Db        437 ATTATCATAAGTGGGTCAATGTTTACTGTGGACAAAGAAATCCTTTCGTCTGCGAGGCTT 496

Qy        362 AGT-CTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              |||  ||||||||  ||||||   ||||||||||||||||||||||||||  |  ||  |  |  |
Db        497 AGTCCTGAAGATGCAGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCACCCCACCTGCG 556

Qy        421 C 421
              |
Db        557 C 557


RESULT 6
AAI71876
ID   AAI71876 standard; DNA; 601 BP.
XX
AC   AAI71876;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin A chain coding sequence.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis; ds.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   P-PSDB; AAM51543.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 1; Page 9; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   encodes the A chain of halyxin.
XX
SQ   Sequence 601 BP; 166 A; 133 C; 163 G; 139 T; 0 other;

  Query Match             48.9%;  Score 266;  DB 22;  Length 601;
  Best Local Similarity   82.4%;  Pred. No. 2.2e-73;
  Matches  305;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||  |||||  | |  ||  ||||||||  ||||  ||||||||  |||
Db        120 GAAGGGCATTGCTACAACATCTTCCATCTATTCAAGACCTGGGCAGAAGCAGAGAGGTTC 179

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||  ||||||||||||  || |  ||||||||||||||||||||||||||  ||||||||||
Db        180 TGCAGGAAGCAGGTGAAGGGCGCGCATCTGGTCTCTATCGAAAGCTCCGAAGAAGCAGAC 239

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  ||||  ||  |   ||  ||||  ||  |||   |   |   |||  ||  ||||||||||
Db        240 TTTGTGGCCCAGCTGGTCTCTGAGAACATGAAGAGATACGGGATCTATATCTGGATCGGA 299

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
```

```
Db        300 CTGAGGGTTCGAGGCAAAAAGAAGCAATGCAGCTCCCAGTGGAGCGATGGCTCCAGCGTC 359

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        360 AGTTATCAGAACTGGATTGAAGCAGAATCCAAAACGTGTCTTGGGCTGCAAAAAGAGACA 419

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Db        420 GAGTTTCGTAAGTGGTTTAATATTTACTGTGGAGAACGAAATCCTTTCGTCTGCGAGGCA 479

Qy        361 TAGTCTGAAG 370

Db        480 TAGAAGCAAG 489
```

RESULT 7
AAQ89309
ID AAQ89309 standard; cDNA; 690 BP.
XX
AC AAQ89309;
XX
DT 25-MAR-2003 (updated)
DT 28-NOV-1995 (first entry)
XX
DE Snake venom antithrombotic oligopeptide cDNA.
XX
KW Antithrombotic peptide; snake venom; platelet binding inhibition;
KW von Willebrand factors; Crotalus horridus horridus; ds.
XX
OS Crotalus horridus horridus.
XX
FH Key Location/Qualifiers
FT CDS 66..515
FT /*tag= a
XX
PN WO9508573-A1.
XX
PD 30-MAR-1995.
XX
PF 21-SEP-1994; 94WO-JP01555.
XX
PR 22-SEP-1993; 93JP-0236975.
XX
PA (AJIN) AJINOMOTO KK.
XX
PI Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
PI Tanaka A, Yamamoto H, Yoshimoto R;
XX
DR WPI; 1995-139559/18.
DR P-PSDB; AAR71981.
XX
PT Single-chain antithrombotic peptide - obtained by cleaving an
PT oligopeptide from snake venom to break inter-chain di:sulphide
PT bonds but preserve intra-chain di:sulphide bonds
XX
PS Example 2; Pages 47-48; 84pp; Japanese.
XX
CC AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
CC specifically from the snake venom oligopeptide AAR71981, encoded by
CC AAQ89309. These peptides have the advantage of avoiding significant
CC thrombocytopenia when administered at the minimun dose, for in vivo
CC inhibition of platelet von Willebrand factor binding.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Query Match 41.6%; Score 226.4; DB 16; Length 690;
Best Local Similarity 70.5%; Pred. No. 7.2e-61;
Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db        168 GATCGGTATTGCTACAAGCCCTTCAAACAAGAGATGACCTGGGCCGATGCAGAGAGGTTC 227

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        228 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTCTCTGTCGAAACCGCCCTAGAAGCATCC 287

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        288 TTTGTGGACAATGTGCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        348 CTGAGGGTTCAAAACAAAGGACAGCCAT----------GCTCCAGCATC 386

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        387 AGTTATGAGAACCTGGTT------GACCCATTTGAATGTTTTATGGTGAGCAGAGACACA 440

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356

Db        441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATTCTTTCATATGCAAGTTC 500

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414

Db        501 ACGCGACCACGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCT 560

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474

Db        561 C--TCCCCACCCCCCACCTTCCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532

Db        619 TCTGTAGCCGGATCTGGGTTTTCTGCTCCAGATGGGTCAGAAGATCCAATAAATTCTGCC 678

Qy        533 TAGCCTGAAAAA 544

Db        679 TACCCAAAAAAA 690
```

RESULT 8
AAC61144
ID AAC61144 standard; DNA; 690 BP.
XX
AC AAC61144;
XX
DT 07-FEB-2001 (first entry)
XX
DE DNA encoding a snake venom derived protein.
XX
KW Subunit peptide production; snake venom; rattlesnake; thrombolytic;
KW von Willebrand's factor; blood platelet-inhibitory activity; ds.
XX
OS Crotalus horridus horridus.
XX
PN WO200059926-A1.
XX
PD 12-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-JP02127.
XX
PR 02-APR-1999; 99JP-0096073.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;
XX
DR WPI; 2000-664985/64.
DR P-PSDB; AAY85628.
XX
PT Producing physiologically-active subunit peptides originating in
PT polymer proteins by denaturation and specific separation, with lower
PT antigenicity but improved solubility and stability, e.g. blood
PT platelet-binding inhibitors -

```
XX
PS   Disclosure; Page 46; 51pp; Japanese.
XX
CC   This invention relates to a method for the production of a subunit
CC   peptide originating from a polymer protein with disulphide bonds within
CC   and between subunits. The method comprises denaturing the protein or its
CC   subunit using a protein denaturing agent in a solution, removing the
CC   agent in the presence of a polyoxyalkyl polyether which reacts with a
CC   thiol group and unwinds the subunit, and separating the polyoxyalkyl
CC   polyether-bound subunit peptide. The method can be used for producing
CC   physiologically-active subunit peptides for polymer proteins e.g. snake
CC   venom-originated dimer peptide with blood platelet-inhibitory activity
CC   on von Willebrand's factor. The peptides produced have platelet-binding
CC   inhibitory, and thrombolytic activity. The present sequence represents
CC   DNA encoding rattlesnake protein used in an example illustrating the
CC   method of the invention.
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

  Query Match             41.6%;  Score 226.4;  DB 21;  Length 690;
  Best Local Similarity   70.5%;  Pred. No. 7.2e-61;
  Matches  389;  Conservative    0;  Mismatches  126;  Indels   37;  Gaps    5;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||  || ||||||||||  |||||||||    | ||||||| | ||||||||||| |||
Db        168 GATCGGTATTGCTACAAGCCCTTCAAACAAGAGATGACCTGGGCCGATGCAGAGAGGTTC 227

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| ||| || ||||||||  ||||| ||||||  | ||  |||||||  |
Db        228 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTCTCTGTCGAAACCGCCCTAGAAGCATCC 287

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  | |   ||  |   |  ||| | | |||    || |    || |||||| |||
Db        288 TTTGTGGACAATGTGCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||||  |||||||||||  | |||  ||                  ||||||||||
Db        348 CTGAGGGTTCAAAACAAAGGACAGCCAT--------------------GCTCCAGCATC 386

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||   | ||      ||  |     | ||| ||  ||||  || ||| |||
Db        387 AGTTATGAGAACCTGGTT------GACCCATTTGAATGTTTTATGGTGAGCAGAGACACA 440

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              || ||| | |||||   ||  || ||||||||||||| || ||||  | ||| |
Db        441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATTCTTTCATATGCAAGTTC 500

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                |  |     | ||||||| ||||  ||||||||||||||||||||||||||||||
Db        501 ACGCGACCACGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCT 560

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              |  ||||| |||   |  || ||||||||||| | | ||||||||||  ||||||||
Db        561 C--TCCCCACCCCCCACCTTCCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||| |||||||  || |||||||  |||||  | ||||   ||||||||||||||
Db        619 TCTGTAGCCGGATCTGGGTTTTCTGCTCCAGATGGGTCAGAAGATCCAATAAATTCTGCC 678

Qy        533 TAGCCTGAAAAA 544
              ||  ||  |||||
Db        679 TACCCAAAAAAA 690


RESULT 9
AAK99834
ID   AAK99834 standard; DNA; 690 BP.
XX
AC   AAK99834;
XX
DT   19-JUL-2002  (first entry)
XX
DE   DNA encoding the antithrombotic wild-type rattlesnake protein.
XX
KW   Thrombolytic; glycoprotein Ib-binding protein; snake venom; rattlesnake;
KW   antithrombotic; glycoprotein Ib; long half life; low antigenicity;
KW   drug efficacy; gene; ds.
XX
OS   Crotalus horridus horridus.
XX
FH   Key             Location/Qualifiers
FT   CDS             66..515
FT                   /*tag=  a
FT                   /product= "Antithrombotic wild-type rattlesnake protein"
XX
PN   EP1195384-A1.
XX
PD   10-APR-2002.
XX
PF   04-OCT-2001; 2001EP-0123277.
XX
PR   04-OCT-2000; 2000JP-0305279.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Fukuchi N,  Kito M,  Kayahara T,  Futaki F,  Ishikawa K,  Suzuki E;
PI   Gondoh K,  Shimba N,  Yamada N;
XX
DR   WPI; 2002-364482/40.
DR   P-PSDB; AAO20974.
XX
PT   Glycoprotein Ib-binding protein, useful for treating thrombosis,
PT   comprises specific mutations in protein originating from snake venom -
XX
PS   Example 1; Page 25; 49pp; English.
XX
CC   The invention relates to a glycoprotein Ib-binding protein, originating
CC   from snake venom, comprising specific mutations and antithrombotic
CC   activity. Glycoprotein Ib-binding protein is used in a drug having
CC   antithrombotic activity. Glycoprotein Ib-binding protein has a high
CC   binding activity to glycoprotein Ib, a long half life/drug efficacy
CC   retention in blood, and low antigenicity. This polynucleotide sequence
CC   represents DNA of the wild-type rattlesnake protein of the invention.
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

  Query Match             41.6%;  Score 226.4;  DB 24;  Length 690;
  Best Local Similarity   70.5%;  Pred. No. 7.2e-61;
  Matches  389;  Conservative    0;  Mismatches  126;  Indels   37;  Gaps    5;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||  || ||||||||||  |||||||||    | ||||||| | ||||||||||| |||
Db        168 GATCGGTATTGCTACAAGCCCTTCAAACAAGAGATGACCTGGGCCGATGCAGAGAGGTTC 227

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| ||| || ||||||||  ||||| ||||||  | ||  |||||||  |
Db        228 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTCTCTGTCGAAACCGCCCTAGAAGCATCC 287

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  | |   ||  |   |  ||| | | |||    || |    || |||||| |||
Db        288 TTTGTGGACAATGTGCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||||  |||||||||||  | |||  ||                  ||||||||||
Db        348 CTGAGGGTTCAAAACAAAGGACAGCCAT--------------------GCTCCAGCATC 386

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||   | ||      ||  |     | ||| ||  ||||  || ||| |||
Db        387 AGTTATGAGAACCTGGTT------GACCCATTTGAATGTTTTATGGTGAGCAGAGACACA 440

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              || ||| | |||||   ||  || ||||||||||||| || ||||  | ||| |
Db        441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATTCTTTCATATGCAAGTTC 500

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                |  |     | ||||||| ||||  ||||||||||||||||||||||||||||||
```

```
Db        501 ACGCGACCACGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCT 560

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              |   ||||| |||    |   || ||||||||||| | | ||||||||||  ||||||||
Db        561 C--TCCCCACCCCCCACCTTCCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||| ||||||||  || ||||||| |||||| | ||||   ||||||||||||||
Db        619 TCTGTAGCCGGATCTGGGTTTTCTGCTCCAGATGGGTCAGAAGATCCAATAAATTCTGCC 678

Qy        533 TAGCCTGAAAAA 544
              || ||  |||||
Db        679 TACCCAAAAAAA 690
```

RESULT 10
AAD32056
ID AAD32056 standard; DNA; 456 BP.
XX
AC AAD32056;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR P-PSDB; AAE20179.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 74; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk3, degenerate nucleic acid.
XX

SQ Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

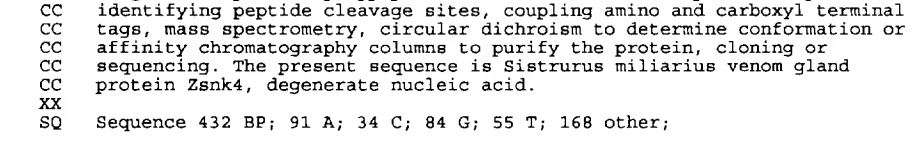
Query Match 34.3%; Score 186.6; DB 24; Length 456;
Best Local Similarity 47.1%; Pred. No. 2.1e-48;
Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   :||:||:||::  || ||:||:||::  ||:|| |||   ||:|| ||::| ||:
Db         97 GAYCARCAYTGYTAYMGNGTNTTYAARCARYTNAARACNTGGGAYGAYGCNGARMGNTTY 156

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||:::  |:||:|   | || || ||::| || :: ||:||::: :: | :||:|| |
Db        157 TGYWSNGARCARGCNGARGGNGGNCAYYTNGTNWSNATHGARWSNWSNGARGARGCNGCN 216

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||:|| |  ||::|  |  |  |:|| :  :     || | : |  |: |:|||||:||
Db        217 TTYGTNGCNCARYTNGTNCCNGARAAYMGNMGNMGNGCNATHYTNTAYATHTGGATHGGN 276

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              :| :| |  ||:   ||:||:||:||:|:::     |:||:: ||:|| :: ::  |
Db        277 YTNMGNGTNCARGGNAARGARAARCARTGYWSNGCNAARTGGWSNGAYGGNWSNWSNGTN 336

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              :: ||:||:||:|||||:||:|  ||::: ||:|  ||:| || | ||   :|  ||
Db        337 WSNTAYGARAAYTGGATHGARGCNGARWSNAARACNTGYYTNGGNYTNCARCARGGNACN 396

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGC 359
                | :||:||:||||  ||: |:||:||:|   |:  : |:|| ||:|| ||:||:||
Db        397 AAYTAYCAYAARTGGGTNAAYATHTAYTGYGGNGARATHAAYCCNTTYGTNTGYGARGC 455
```

RESULT 11
AAD32058
ID AAD32058 standard; DNA; 432 BP.
XX
AC AAD32058;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 76; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet

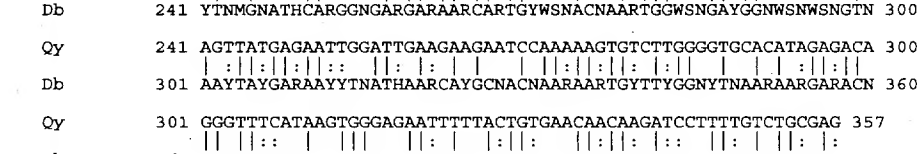
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk4, degenerate nucleic acid.
XX
SQ Sequence 432 BP; 91 A; 34 C; 84 G; 55 T; 168 other;

Query Match 32.2%; Score 175; DB 24; Length 432;
Best Local Similarity 45.7%; Pred. No. 9.3e-45;
Matches 163; Conservative 85; Mismatches 109; Indels 0; Gaps 0;



RESULT 12
AAD32060
ID AAD32060 standard; DNA; 474 BP.
XX
AC AAD32060;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk5, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 78; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk5, degenerate nucleic acid.
XX
SQ Sequence 474 BP; 89 A; 31 C; 84 G; 74 T; 196 other;

Query Match 30.5%; Score 166; DB 24; Length 474;
Best Local Similarity 43.4%; Pred. No. 6.7e-42;
Matches 155; Conservative 88; Mismatches 114; Indels 0; Gaps 0;



RESULT 13
AAI71877
ID AAI71877 standard; DNA; 633 BP.
XX
AC AAI71877;
XX
DT 10-JAN-2002 (first entry)
XX
DE Snake venom blood anticoagulant halyxin B chain coding sequence.
XX
KW Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW thrombosis; ds.

XX
OS Agkistrodon halys.
XX
PN KR2001049671-A.
XX
PD 15-JUN-2001.
XX
PF 29-JUN-2000; 2000KR-0036591.
XX
PR 29-JUN-1999; 99KR-0025105.
XX
PA (BIOB-) BIOBUD CO LTD.
XX
PI Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
XX
DR WPI; 2001-637330/73.
DR P-PSDB; AAM51544.
XX
PT Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS Claim 1; Page 11; 21pp; Korean.
XX
CC The invention relates to halyxin, a novel protein with very strong
CC blood anticoagulation activity. The protein was separated from snake
CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC be used in the treatment of thrombogenesis. The present sequence
CC encodes the B chain of halyxin.
XX
SQ Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;

Query Match 26.5%; Score 144; DB 22; Length 633;
Best Local Similarity 63.4%; Pred. No. 6.7e-35;
Matches 350; Conservative 0; Mismatches 140; Indels 62; Gaps 6;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||    || ||  ||   || |  ||| | ||||||||||  |||
Db        120 GAAGGGCATTGCTACAAACCATTTAATGAACAAAAAAATTGGGCCGATGCAGAGAATTTC 179

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||||  | |||    |  || || ||||||||||  || | ||| | | |||||||||
Db        180 TGCACACAACAGCACACAGGCGGACATCTGGTCTCCTTCCACAGCACTGAAGAAGCAGAT 239

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  | || ||      ||      ||| |  | | |       | |||||||  |||
Db        240 TTTGTGGTCAAGCTGGCCTTCCA------AAATTTTGGCCACGGTATTTTCTGGATGGGA 293

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||       ||        || |||||||||     | ||||||  ||   |||   ||
Db        294 CTGAG------CAATGTCTGGAATCAATGCAGCTGGCAATGGAGCAGTGCTGCCAAGCTC 347

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |  || ||    |||  |||||||          |  |  | || |  |  || ||
Db        348 AAATACGAAGCCTGGGCTGAAGAA----------TCTTATTGTGTCTATTTCAAGTCA 395

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCG----- 355
                | | || |||  || |     |||   |    | |    ||| |||||||
Db        396 ACTAATAACAAATGGAGGAGTAGAGCCTGCAGAATGGAGGCATATTTCGTCTGCGAGTTC 455

Qy        356 -AGGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
               ||||||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db        456 CAGGCATAGTTTGAAGATCCAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACAT 515

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              || ||||                          |||||||  |||||| ||||||||
Db        516 CACCCCC----------------------------CACCCTTCGCTCAATGGATGCTC 545

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||||||||||||| ||||||||||||||||||| ||||   ||||||||||||||
Db        546 TCTGTAGCTGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCC 605

Qy        533 TAGCCTGAAAAA 544
              |||| |||||||
Db        606 TAGCATGAAAAA 617
```

RESULT 14
AAQ89308
ID AAQ89308 standard; cDNA; 272 BP.
XX
AC AAQ89308;
XX
DT 25-MAR-2003 (updated)
DT 28-NOV-1995 (first entry)
XX
DE Snake venom antithrombotic oligopeptide cDNA.
XX
KW Antithrombotic peptide; snake venom; platelet binding inhibition;
KW von Willebrand factors; Crotalus horridus horridus; ds.
XX
OS Crotalus horridus horridus.
XX
PN WO9508573-A1.
XX
PD 30-MAR-1995.
XX
PF 21-SEP-1994; 94WO-JP01555.
XX
PR 22-SEP-1993; 93JP-0236975.
XX
PA (AJIN) AJINOMOTO KK.
XX
PI Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
PI Tanaka A, Yamamoto H, Yoshimoto R;
XX
DR WPI; 1995-139559/18.
XX
PT Single-chain antithrombotic peptide - obtained by cleaving an
PT oligopeptide from snake venom to break inter-chain di:sulphide
PT bonds but preserve intra-chain di:sulphide bonds
XX
PS Example 2; Page 46; 84pp; Japanese.
XX
CC AAQ89308 is a snake venom cDNA used in the prepn. of AAQ89309, which
CC encodes AAR71981 a snake venom oligopeptide. From this oligopeptide
CC claimed antithrobtic peptides were derived, the peptides had the
CC advantage of avoiding significant thrombocytopenia when administered
CC at the minimun dose, for in vivo inhibition of platelet von Willebrand
CC factor binding.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 272 BP; 72 A; 60 C; 77 G; 63 T; 0 other;

Query Match 18.9%; Score 102.6; DB 16; Length 272;
Best Local Similarity 72.9%; Pred. No. 5e-22;
Matches 132; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
Qy         34 AAGACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 93
              | ||| ||| | |||||||||||| |||||| || |||||| ||| || ||||||||  ||
Db          7 ATGACTTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 66

Qy         94 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAG 153
              ||| |||||| | ||  |||||||  |||||||| | |  || |   | ||| | | ||
Db         67 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 126

Qy        154 TCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 213
              |    || |    || ||||||| |||||||||| |||||||||| | ||| ||||  |
Db        127 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGCTCC 186

Qy        214 A 214
              |
Db        187 A 187
```

RESULT 15
AAC61143
ID AAC61143 standard; DNA; 272 BP.

XX
AC AAC61143;
XX
DT 07-FEB-2001 (first entry)
XX
DE Rattlesnake venom related DNA SEQ ID 4.
XX
KW Subunit peptide production; snake venom; rattlesnake; thrombolytic; ds;
KW von Willebrand's factor; blood platelet-inhibitory activity.
XX
OS Synthetic.
XX
PN WO200059926-A1.
XX
PD 12-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-JP02127.
XX
PR 02-APR-1999; 99JP-0096073.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;
XX
DR WPI; 2000-664985/64.
XX
PT Producing physiologically-active subunit peptides originating in
PT polymer proteins by denaturation and specific separation, with lower
PT antigenicity but improved solubility and stability, e.g. blood
PT platelet-binding inhibitors -
XX
PS Example 1; Page 45; 51pp; Japanese.
XX
CC This invention relates to a method for the production of a subunit
CC peptide originating from a polymer protein with disulphide bonds within
CC and between subunits. The method comprises denaturing the protein or its
CC subunit using a protein denaturing agent in a solution, removing the
CC agent in the presence of a polyoxyalkyl polyether which reacts with a
CC thiol group and unwinds the subunit, and separating the polyoxyalkyl
CC polyether-bound subunit peptide. The method can be used for producing
CC physiologically-active subunit peptides for polymer proteins e.g. snake
CC venom-originated dimer peptide with blood platelet-inhibitory activity
CC on von Willebrand's factor. The peptides produced have platelet-binding
CC inhibitory, and thrombolytic activity. The present sequence represents a
CC DNA related to a rattlesnake protein used in an example illustrating the
CC method of the invention.
XX
SQ Sequence 272 BP; 72 A; 60 C; 77 G; 63 T; 0 other;

Query Match 18.9%; Score 102.6; DB 21; Length 272;
Best Local Similarity 72.9%; Pred. No. 5e-22;
Matches 132; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
Qy         34 AAGACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 93
              | ||| ||| | |||||||||||| |||||| || |||||| ||| || ||||||||  ||
Db          7 ATGACTTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 66

Qy         94 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAG 153
              ||| |||||| | ||  ||||||| |||||||| | |  || |   |  ||| | | ||
Db         67 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 126

Qy        154 TCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 213
              |    || |    || ||||||| |||||||||| |||||||||| | ||| ||||  |
Db        127 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGCTCC 186

Qy        214 A 214
              |
Db        187 A 187
```

Search completed: December 8, 2003, 11:08:47
Job time : 323 secs

This Page Blank (uspto)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:25:43 ; Search time 79 Seconds
(without alignments)
3039.398 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTACAAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	544	100.0	544	4	US-09-058-740-1	Sequence 1, Appli
	2	226.4	41.6	690	2	US-08-612-840A-7	Sequence 7, Appli
	3	102.6	18.9	272	2	US-08-612-840A-6	Sequence 6, Appli
c	4	58.4	10.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	58	10.7	454	4	US-09-058-740-12	Sequence 12, Appl
	6	34.6	6.4	2949	4	US-09-412-554A-3	Sequence 3, Appli
	7	34.2	6.3	590	2	US-08-454-557C-32	Sequence 32, Appl
	8	34.2	6.3	590	2	US-08-340-426D-32	Sequence 32, Appl
	9	34.2	6.3	590	2	US-08-450-673C-32	Sequence 32, Appl
	10	34.2	6.3	590	5	PCT-US95-17111A-32	Sequence 32, Appl
	11	32.6	6.0	777	3	US-09-146-969-1	Sequence 1, Appli
	12	32.4	6.0	614	2	US-08-729-103-2	Sequence 2, Appli
	13	32.4	6.0	1114	2	US-08-468-413-1	Sequence 1, Appli
	14	32.4	6.0	1114	3	US-09-162-508-1	Sequence 1, Appli
	15	32.4	6.0	1114	5	PCT-US95-07169-1	Sequence 1, Appli
c	16	32.2	5.9	562	4	US-09-669-751-15	Sequence 15, Appl
	17	32	5.9	1201	4	US-09-461-325-42	Sequence 42, Appl
	18	31.6	5.8	531	1	US-08-340-539A-5	Sequence 5, Appli
	19	31.6	5.8	531	2	US-08-461-592B-5	Sequence 5, Appli
	20	31.6	5.8	1829	6	5514582-1	Patent No. 5514582
	21	31.6	5.8	2259	2	US-08-513-278-1	Sequence 1, Appli
	22	31.6	5.8	2330	1	US-08-481-803-1	Sequence 1, Appli
	23	31.6	5.8	2330	1	US-08-215-366A-1	Sequence 1, Appli
	24	31.6	5.8	2330	1	US-08-340-539A-1	Sequence 1, Appli
	25	31.6	5.8	2330	2	US-08-461-592B-1	Sequence 1, Appli
	26	31.6	5.8	2385	4	US-09-495-050A-292	Sequence 292, App
	27	31.6	5.8	5169	4	US-09-194-612A-2	Sequence 2, Appli
	28	31.4	5.8	8906	2	US-08-826-267-1	Sequence 1, Appli
	29	31.2	5.7	46819	4	US-09-453-702B-72	Sequence 72, Appl
c	30	30.8	5.7	5261	1	US-08-045-806-3	Sequence 3, Appli
c	31	30.8	5.7	5261	1	US-08-366-051B-3	Sequence 3, Appli
c	32	30.6	5.6	280	6	5198345-11	Patent No. 5198345
	33	30.6	5.6	289	3	US-09-007-005-17	Sequence 17, Appl
	34	30.6	5.6	289	3	US-09-244-796-17	Sequence 17, Appl
c	35	30.6	5.6	289	6	5198345-18	Patent No. 5198345
	36	30.4	5.6	747	2	US-08-401-530A-1	Sequence 1, Appli
	37	30.4	5.6	2325	2	US-08-714-677-3	Sequence 3, Appli
	38	30.4	5.6	2325	2	US-08-393-540-3	Sequence 3, Appli
	39	30.4	5.6	2325	2	US-08-714-537-3	Sequence 3, Appli
	40	30.4	5.6	3073	2	US-08-714-677-11	Sequence 11, Appl
	41	30.4	5.6	3073	2	US-08-393-540-11	Sequence 11, Appl
	42	30.4	5.6	3073	2	US-08-714-537-11	Sequence 11, Appl
	43	30.2	5.6	762	2	US-08-822-261-2	Sequence 2, Appli
	44	30.2	5.6	762	4	US-09-226-852-2	Sequence 2, Appli
c	45	30	5.5	406	4	US-09-556-877-50	Sequence 50, Appl

ALIGNMENTS

```
RESULT 1
US-09-058-740-1
; Sequence 1, Application US/09058740
; Patent No. 6489451
;   GENERAL INFORMATION:
;        APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                   Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                   Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                   Jin-Guo Ding, Fang Rong, Yan Liu and
;                   Hui-Ran Chen
;        TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                            VENOM OF AGKISTRODON ACUTUS
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/058,740
;             FILING DATE: 10-Apr-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Chen, Anthony C.
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 544 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-058-740-1
```

```
  Query Match             100.0%;  Score 544;  DB 4;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-164;
  Matches  544;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540

Qy        541 AAAA 544
              ||||
Db        541 AAAA 544
```

```
RESULT 2
US-08-612-840A-7
; Sequence 7, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  690 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Crotalus horridus horridus
;      STRAIN:
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  66..512
US-08-612-840A-7
```

```
  Query Match             41.6%;  Score 226.4;  DB 2;  Length 690;
  Best Local Similarity   70.5%;  Pred. No. 8.8e-63;
  Matches  389;  Conservative    0;  Mismatches  126;  Indels   37;  Gaps    5;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||  || |||||||||||  |||||||||    | |||||| | ||||||||||| |||
Db        168 GATCGGTATTGCTACAAGCCCTTCAAACAAGAGATGACCTGGGCCGATGCAGAGAGGTTC 227

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| ||| || ||||||||  ||||| |||||| | ||  |||||||  |
Db        228 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTCTCTGTCGAAACCGCCCTAGAAGCATCC 287

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| | |  || |   |  ||| | | |||    || |    || |||||| |||
Db        288 TTTGTGGACAATGTGCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||||  |||||||||||  | |||  ||                    ||||||||||
Db        348 CTGAGGGTTCAAAACAAAGGACAGCCAT--------------------GCTCCAGCATC 386

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||    | ||      ||  |     | ||| ||  ||||  || ||| |||
Db        387 AGTTATGAGAACCTGGTT------GACCCATTTGAATGTTTTATGGTGAGCAGAGACACA 440

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              || ||| | |||||   ||  || ||||||||||||| || ||||  | ||| |
Db        441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATTCTTTCATATGCAAGTTC 500

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                |  |    | |||||| |||| ||||||||||||||||||||||||||||||
Db        501 ACGCGACCACGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCT 560

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474
              |   ||||| |||    |  || ||||||||||| | | ||||||||||  ||||||||
Db        561 C--TCCCCACCCCCCACCTTCCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532
              |||||||| ||||||||  || ||||||| |||||| | ||||   ||||||||||||||
Db        619 TCTGTAGCCGGATCTGGGTTTTCTGCTCCAGATGGGTCAGAAGATCCAATAAATTCTGCC 678

Qy        533 TAGCCTGAAAAA 544
              || ||  |||||
```

```
Db        679 TACCCAAAAAAA 690


RESULT 3
US-08-612-840A-6
; Sequence 6, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  272 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Crotalus horridus horridus
;      STRAIN:
US-08-612-840A-6

  Query Match             18.9%;  Score 102.6;  DB 2;  Length 272;
  Best Local Similarity   72.9%;  Pred. No. 2.6e-23;
  Matches  132;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy         34 AAGACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 93
              | ||| ||| | |||||||||||| |||||| || |||||| ||| || ||||||||  ||
Db          7 ATGACTTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 66

Qy         94 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAG 153
              ||| |||||| | ||  ||||||| |||||||| | |  || |   | ||| | | ||
Db         67 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 126

Qy        154 TCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 213
              |    || |    || ||||||| |||||||||| |||||||||| | ||| ||||  |
Db        127 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGCTCC 186

Qy        214 A 214
              |
Db        187 A 187


RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             10.7%;  Score 58.4;  DB 1;  Length 7218;
  Best Local Similarity   6.2%;  Pred. No. 2e-08;
  Matches   26;  Conservative  222;  Mismatches  168;  Indels    0;  Gaps    0;

Qy         11 GCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGAAGC 70
              | || |||   ||   ||::  ::::  ::: ::: : ::::::     : : ::::
Db       1452 GATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1393

Qy         71 AGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCC 130
              ::: ::: :::::: :   ::    :  :::::    ::::::: :::    : :::
Db       1392 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1333

Qy        131 AGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTC 190
              ::  ::  :  :::::: :::  ::  ::::   : :   :::  :::  :::::
Db       1332 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1273

Qy        191 AAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGA 250
              :::: ::::::::: :: : :: : :::: :::: :: ::    :: :  ::  : ::::
```

```
Db       1272 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1213

Qy        251 ATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAGGGTTTCATA 310

Db       1212 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1153

Qy        311 AGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCATAGTCTGAAG 370

Db       1152 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1093

Qy        371 ATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCCCCAACC 426

Db       1092 RRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGC 1037
```

RESULT 5
US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,740
; FILING DATE: 10-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...440
; OTHER INFORMATION: "N" stands for any base.
; "Xaa" stands for any amino acid.
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-058-740-12

Query Match 10.7%; Score 58; DB 4; Length 454;
Best Local Similarity 77.8%; Pred. No. 6.7e-09;
Matches 70; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         99 GAAGGGCATTGCTACAAGCCCTTCGATGAACCTAAGACCTGGGCAGATGCAGAGAAATTC 158

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTG 90

Db        159 TGCACACAACAACACAAAGGCAGCCATCTG 188
```

RESULT 6
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate oligonucleotide sequence encoding the
; OTHER INFORMATION: zfsta2 polypeptide of SEQ ID NO:2.
; NAME/KEY: variation
; LOCATION: (1)...(2949)
; OTHER INFORMATION: Each N is independently any nucleotide.
US-09-412-554A-3

Query Match 6.4%; Score 34.6; DB 4; Length 2949;
Best Local Similarity 29.8%; Pred. No. 0.53;
Matches 68; Conservative 43; Mismatches 117; Indels 0; Gaps 0;

```
Qy         67 AAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTG 126

Db        292 AARCCNGTNTGYGGNWSNGAYGGNGARTTYTAYGARAAYCAYTGYGARGTNCAYMGNGCN 351

Qy        127 GGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGG 186

Db        352 GCNTGYYTNAARAARCARAARATHACNATHGTNCAYAAYGARGAYTGYTTYTTYAARGGN 411

Qy        187 GCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTAT 246

Db        412 GAYAARTGYAARACNACNGARTAYWSNAARATGAARAAYATGYTNYTNGAYYTNCARAAY 471

Qy        247 GAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATA 294

Db        472 CARAARTAYATHATGCARGARAAYGARAAYCCNAAYGGNGAYGAYATH 519
```

RESULT 7
US-08-454-557C-32
; Sequence 32, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.

```
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/454,557C
;       FILING DATE:  30-MAY-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ludwig, Steven R.
;       REGISTRATION NUMBER:  36,203
;       REFERENCE/DOCKET NUMBER:  0609.3840003
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  32:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  590 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
US-08-454-557C-32

  Query Match             6.3%;  Score 34.2;  DB 2;  Length 590;
  Best Local Similarity   52.1%;  Pred. No. 0.32;
  Matches  190;  Conservative    0;  Mismatches  148;  Indels   27;  Gaps    4;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              | ||||||  |  |      || |   | ||||||||   ||||||||   ||  |||
Db         16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
              ||  |  ||||   |||  | ||||| |||  |    | |   |  || |  | ||||||||
Db         73 AGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAGGGTGCCTTTGTGG 132

Qy        128 GCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGG 187
              |    ||||   | || |||      | |   |  || ||||||||| || ||
Db        133 CCTCACTGAT---TAAGGAGAGTGGCACTGATGACTTCAATGTCTGGATTGGCCT----- 184

Qy        188 CTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATG 247
                ||  ||    ||||| |  || ||     | ||||||  ||| |||    ||   ||
Db        185 -CCATGACCCCAAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGGTCTCCTACA 243

Qy        248 AGAATTGGATTGAAGAAGAATCCAAAA---------------AGTGTCTTGGGGTGCACA 292
              ||   |||       | ||   | |  |               | ||| |  |  ||  |
Db        244 AGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGTGTGAGCCTGACCT 303

Qy        293 TAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCT 352
              |  |||||| || || || ||| || || |    ||||||  |  |     | |||||||
Db        304 CAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAGTTCTCCTTTGTCT 363

Qy        353 GCGAG 357
              || ||
Db        364 GCAAG 368


RESULT 8
US-08-340-426D-32
; Sequence 32, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-32

  Query Match             6.3%;  Score 34.2;  DB 2;  Length 590;
  Best Local Similarity   52.1%;  Pred. No. 0.32;
  Matches  190;  Conservative    0;  Mismatches  148;  Indels   27;  Gaps    4;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              | ||||||  |  |      || |   | ||||||||   ||||||||   ||  |||
Db         16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
              ||  |  ||||   |||  | ||||| |||  |    | |   |  || |  | ||||||||
Db         73 AGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAGGGTGCCTTTGTGG 132

Qy        128 GCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGG 187
              |    ||||   | || |||      | |   |  || ||||||||| || ||
Db        133 CCTCACTGAT---TAAGGAGAGTGGCACTGATGACTTCAATGTCTGGATTGGCCT----- 184

Qy        188 CTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATG 247
                ||  ||    ||||| |  || ||     | ||||||  ||| |||    ||   ||
Db        185 -CCATGACCCCAAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGGTCTCCTACA 243

Qy        248 AGAATTGGATTGAAGAAGAATCCAAAA---------------AGTGTCTTGGGGTGCACA 292
              ||   |||       | ||   | |  |               | ||| |  |  ||  |
Db        244 AGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGTGTGAGCCTGACCT 303

Qy        293 TAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCT 352
              |  |||||| || || || ||| || || |    ||||||  |  |     | |||||||
Db        304 CAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAGTTCTCCTTTGTCT 363

Qy        353 GCGAG 357
              || ||
Db        364 GCAAG 368


RESULT 9
US-08-450-673C-32
; Sequence 32, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
```

```
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-32

  Query Match             6.3%;  Score 34.2;  DB 2;  Length 590;
  Best Local Similarity   52.1%;  Pred. No. 0.32;
  Matches  190;  Conservative    0;  Mismatches  148;  Indels   27;  Gaps    4;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              | ||||||  |  |     ||  |   | ||||||||   ||||||||   ||  |||
Db         16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
              ||  |  ||||   |||  | ||||| ||| |    | |   |  || |  | ||||||||
Db         73 AGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAGGGTGCCTTTGTGG 132

Qy        128 GCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGG 187
               |    ||||   | || |||       | |   |  || ||||||||| || ||
Db        133 CCTCACTGAT---TAAGGAGAGTGGCACTGATGACTTCAATGTCTGGATTGGCCT----- 184

Qy        188 CTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATG 247
               ||  ||    ||||| |  || ||    | ||||||  ||| |||   ||   ||
Db        185 -CCATGACCCCAAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGGTCTCCTACA 243

Qy        248 AGAATTGGATTGAAGAAGAATCCAAAA--------------AGTGTCTTGGGGTGCACA 292
              ||   |||       | ||   | |  |             | ||| |  |  ||  |
Db        244 AGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGTGTGAGCCTGACCT 303

Qy        293 TAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCT 352
               |   ||||| || || || ||| || || |    ||||| | |      | |||||||
Db        304 CAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAGTTCTCCTTTGTCT 363

Qy        353 GCGAG 357
              || ||
Db        364 GCAAG 368


RESULT 10
PCT-US95-17111A-32
; Sequence 32, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
```

```
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-32

  Query Match             6.3%;  Score 34.2;  DB 5;  Length 590;
  Best Local Similarity   52.1%;  Pred. No. 0.32;
  Matches  190;  Conservative    0;  Mismatches  148;  Indels   27;  Gaps    4;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              | ||||||  |  |     ||  |   | ||||||||   ||||||||   ||  |||
Db         16 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 72

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
              ||  |  ||||   |||  | ||||| ||| |    | |   |  || |  | ||||||||
Db         73 AGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAGGGTGCCTTTGTGG 132

Qy        128 GCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGG 187
               |    ||||   | || |||       | |   |  || ||||||||| || ||
Db        133 CCTCACTGAT---TAAGGAGAGTGGCACTGATGACTTCAATGTCTGGATTGGCCT----- 184

Qy        188 CTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATG 247
               ||  ||    ||||| |  || ||    | ||||||  ||| |||   ||   ||
Db        185 -CCATGACCCCAAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGGTCTCCTACA 243

Qy        248 AGAATTGGATTGAAGAAGAATCCAAAA--------------AGTGTCTTGGGGTGCACA 292
              ||   |||       | ||   | |  |             | ||| |  |  ||  |
Db        244 AGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGTGTGAGCCTGACCT 303

Qy        293 TAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCT 352
               |   ||||| || || || ||| || || |    ||||| | |      | |||||||
Db        304 CAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAGTTCTCCTTTGTCT 363

Qy        353 GCGAG 357
              || ||
Db        364 GCAAG 368


RESULT 11
US-09-146-969-1
; Sequence 1, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
;  APPLICANT: Dieckgraefe, Brian K.
;  TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
;  FILE REFERENCE: 04255.75314
;  CURRENT APPLICATION NUMBER: US/09/146,969
;  CURRENT FILING DATE: 1998-09-04
```

```
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 777
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-146-969-1

  Query Match             6.0%;  Score 32.6;  DB 3;  Length 777;
  Best Local Similarity   51.8%;  Pred. No. 1.2;
  Matches  189;  Conservative    0;  Mismatches  149;  Indels   27;  Gaps    4;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              | |||||| |  |      || |   | ||||||||   ||||||||    ||  |||
Db        219 ACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGATCTCTATTGC---C 275

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
              || |  ||||   |||  | ||||| ||| |    | |   |  || |  | ||||||||
Db        276 AGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAGGGTGCCTTTGTGG 335

Qy        128 GCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGG 187
              |    ||||   | || |||      | |   |  || ||||||||| || ||
Db        336 CCTCACTGAT---TAAGGAGAGTGGCACTGATGACTTCAATGTCTGGATTGGCCT----- 387

Qy        188 CTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATG 247
              ||  ||   ||||| |  || ||    | ||||||  ||| |||   ||  ||
Db        388 -CCATGACCCCAAAAAGAACCGCCGCTGGCACTGGAGCAGTGGGTCCCTGGTCTCCTACA 446

Qy        248 AGAATTGGATTGAAGAAGAATCCAAAA---------------AGTGTCTTGGGGTGCACA 292
              ||  |||     | ||  | |  |                 | ||| |  |  ||  |
Db        447 AGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGTGTGAGCCTGACCT 506

Qy        293 TAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCT 352
              |   ||||| || || || ||| || || |    ||||||  |  |      | ||||| |
Db        507 CAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAGTTCTCCTTTGTAT 566

Qy        353 GCGAG 357
              || ||
Db        567 GCAAG 571


RESULT 12
US-08-729-103-2
; Sequence 2, Application US/08729103
; Patent No. 5837841
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN REG PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/729,103
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0138 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  614 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      LIBRARY:  COLNFET02
;      CLONE:  1310334
US-08-729-103-2

  Query Match             6.0%;  Score 32.4;  DB 2;  Length 614;
  Best Local Similarity   53.1%;  Pred. No. 1.2;
  Matches   69;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              |||||||       |||||   |    | || |||| | ||||| |||  |   ||   |
Db        254 ATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTGTCAGT 313

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
                | |  ||||| |  || ||||  ||||||   || |     |||||   |    | |
Db        314 CTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACCATAG 373

Qy        128 GCCAGTTGAT 137
              |||  ||
Db        374 CAGAGTACAT 383


RESULT 13
US-08-468-413-1
; Sequence 1, Application US/08468413
; Patent No. 5861494
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Colon Specific Gene
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,413
;      FILING DATE:  06 JUN 95
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-447
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1114 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
```

```
US-08-468-413-1

  Query Match             6.0%;  Score 32.4;  DB 2;  Length 1114;
  Best Local Similarity   53.1%;  Pred. No. 1.7;
  Matches   69;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              |||||||       |||||    |     | || |||| | ||||| |||  |    ||   |
Db        229 ATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTGTCAGT 288

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
                | |  ||||| |  || ||||  ||||||   || |     |||||   |    | |
Db        289 CTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACCATAG 348

Qy        128 GCCAGTTGAT 137
                |||  ||
Db        349 CAGAGTACAT 358


RESULT 14
US-09-162-508-1
; Sequence 1, Application US/09162508
; Patent No. 6080722
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Colon Specific Gene
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,508
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/468,413
;      FILING DATE:  06 JUN 95
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-447
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1114 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
US-09-162-508-1

  Query Match             6.0%;  Score 32.4;  DB 3;  Length 1114;
  Best Local Similarity   53.1%;  Pred. No. 1.7;
  Matches   69;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              |||||||       |||||    |     | || |||| | ||||| |||  |    ||   |
Db        229 ATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTGTCAGT 288

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
                | |  ||||| |  || ||||  ||||||   || |     |||||   |    | |
Db        289 CTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACCATAG 348

Qy        128 GCCAGTTGAT 137
                |||  ||
Db        349 CAGAGTACAT 358


RESULT 15
PCT-US95-07169-1
; Sequence 1, Application PC/TUS9507169
;  GENERAL INFORMATION:
;    APPLICANT:  LI, ET AL.
;    TITLE OF INVENTION:  Human Colon Specific Gene
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN,
;      ADDRESSEE:  CECCHI, STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NEW JERSEY
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH DISKETTE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07169
;      FILING DATE:  06 JUN 95
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FERRARO, GREGORY D.
;      REGISTRATION NUMBER:  36,134
;      REFERENCE/DOCKET NUMBER:  325800-389
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1114 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  cDNA
PCT-US95-07169-1

  Query Match             6.0%;  Score 32.4;  DB 5;  Length 1114;
  Best Local Similarity   53.1%;  Pred. No. 1.7;
  Matches   69;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy          8 ATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTGCACGA 67
              |||||||       |||||    |     | || |||| | ||||| |||  |    ||   |
Db        229 ATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTGTCAGT 288

Qy         68 AGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGG 127
                | |  ||||| |  || ||||  ||||||   || |     |||||   |    | |
Db        289 CTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACCATAG 348

Qy        128 GCCAGTTGAT 137
                |||  ||
Db        349 CAGAGTACAT 358


Search completed: December  8, 2003, 12:41:19
Job time : 81 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 12:02:04 ; Search time 346 Seconds
(without alignments)
5225.543 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTACAAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	544	100.0	544	11	US-09-938-114-1	Sequence 1, Appli
2	306.2	56.3	721	10	US-09-929-230-4	Sequence 4, Appli
3	306.2	56.3	721	13	US-10-226-420-4	Sequence 4, Appli
4	288	52.9	725	10	US-09-929-230-10	Sequence 10, Appl
5	288	52.9	725	13	US-10-226-420-10	Sequence 10, Appl
6	285.8	52.5	580	10	US-09-929-230-7	Sequence 7, Appli
7	285.8	52.5	580	13	US-10-226-420-7	Sequence 7, Appli
8	226.4	41.6	690	10	US-09-969-763-2	Sequence 2, Appli
9	186.6	34.3	456	10	US-09-929-230-6	Sequence 6, Appli
10	186.6	34.3	456	13	US-10-226-420-6	Sequence 6, Appli
11	175	32.2	432	10	US-09-929-230-9	Sequence 9, Appli
12	175	32.2	432	13	US-10-226-420-9	Sequence 9, Appli
13	166	30.5	474	10	US-09-929-230-12	Sequence 12, Appl
14	166	30.5	474	13	US-10-226-420-12	Sequence 12, Appl
15	82.4	15.1	483	10	US-09-929-230-1	Sequence 1, Appli
16	82.4	15.1	483	13	US-10-226-420-1	Sequence 1, Appli
17	64.2	11.8	453	10	US-09-929-230-3	Sequence 3, Appli
18	64.2	11.8	453	13	US-10-226-420-3	Sequence 3, Appli
19	36.8	6.8	1597	9	US-09-726-643-30	Sequence 30, Appl
20	36.8	6.8	1597	14	US-10-042-141-30	Sequence 30, Appl
21	36.4	6.7	498	10	US-09-920-300A-1755	Sequence 1755, Ap
22	36.4	6.7	498	13	US-10-099-926-1755	Sequence 1755, Ap
23	36.4	6.7	498	14	US-10-033-528-1755	Sequence 1755, Ap
24	36	6.6	492	9	US-09-801-438-3	Sequence 3, Appli
c 25	35.6	6.5	1223197	13	US-10-027-632-179264	Sequence 179264,
c 26	35.6	6.5	1223197	14	US-10-027-632-179264	Sequence 179264,
27	34.8	6.4	461	10	US-09-920-300A-1761	Sequence 1761, Ap
28	34.8	6.4	461	13	US-10-099-926-1761	Sequence 1761, Ap
29	34.8	6.4	461	14	US-10-033-528-1761	Sequence 1761, Ap
c 30	34.8	6.4	509	15	US-10-060-036-2993	Sequence 2993, Ap
31	34.6	6.4	2949	15	US-10-013-136-3	Sequence 3, Appli
32	34.2	6.3	493	10	US-09-920-300A-1786	Sequence 1786, Ap
33	34.2	6.3	493	13	US-10-099-926-1786	Sequence 1786, Ap
34	34.2	6.3	493	14	US-10-033-528-1786	Sequence 1786, Ap
35	34.2	6.3	800	13	US-09-997-003-11	Sequence 11, Appl
36	34.2	6.3	843	9	US-09-925-301-340	Sequence 340, App
37	34.2	6.3	843	13	US-09-997-003-24	Sequence 24, Appl
38	33.8	6.2	585	10	US-09-920-300A-1778	Sequence 1778, Ap
39	33.8	6.2	585	13	US-10-099-926-1778	Sequence 1778, Ap
40	33.8	6.2	585	14	US-10-033-528-1778	Sequence 1778, Ap
41	33.8	6.2	593	10	US-09-920-300A-1756	Sequence 1756, Ap
42	33.8	6.2	593	13	US-10-099-926-1756	Sequence 1756, Ap
43	33.8	6.2	593	14	US-10-033-528-1756	Sequence 1756, Ap
44	33.8	6.2	671	15	US-10-184-644-346	Sequence 346, App
45	33.8	6.2	671	15	US-10-184-634-346	Sequence 346, App

ALIGNMENTS

```
RESULT 1
US-09-938-114-1
; Sequence 1, Application US/09938114
; Publication No. US20030022350A1
;  GENERAL INFORMATION:
;       APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                  Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                  Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                  Jin-Guo Ding, Fang Rong, Yan Liu and
;                  Hui-Ran Chen
;       TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                           VENOM OF AGKISTRODON ACUTUS
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/938,114
;            FILING DATE: 23-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/058,740
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Chen, Anthony C.
;            REGISTRATION NUMBER: 38,673
;            REFERENCE/DOCKET NUMBER: 233/298
```

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 544 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-938-114-1

  Query Match            100.0%;  Score 544;  DB 11;  Length 544;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-175;
  Matches  544;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         61 TGCACGAAGCAGGTGAACGGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCCACCCCATCCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCCACCCCATCCC 420

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTA 480

Qy        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAGTCAATAAATTCTGCCTAGCCTGA 540

Qy        541 AAAA 544
              ||||
Db        541 AAAA 544


RESULT 2
US-09-929-230-4
; Sequence 4, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 721
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)...(546)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk3
US-09-929-230-4

  Query Match            56.3%;  Score 306.2;  DB 10;  Length 721;
  Best Local Similarity  78.8%;  Pred. No. 5e-94;
  Matches  365;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   |||||||||||| ||||||||||||||    ||||| |||   |||||||||||| |||
Db        187 GATCAGCATTGCTACAGGGTCTTCAAACAACTCAAGACGTGGGACGATGCAGAGAGGTTC 246

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| | | || |||||||| ||||||||||||||||||| ||||||||| |
Db        247 TGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTCTCTATCGAAAGCTCCGAAGAAGCAGCC 306

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || |  || |||| |  | |  |||||   || || ||||||||||||
Db        307 TTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGGAGAGCCATTCTCTATATCTGGATCGGA 366

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| ||||  |||||| ||||||||||||    ||||||||||||||||||| ||
Db        367 CTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGCGCGAAGTGGAGCGATGGCTCCAGCGTC 426

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              |||||||||||| |||||||||| |||||||||| |||||||| ||||   ||  |||
Db        427 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACATGTCTTGGGCTGCAACAAGGCACA 486

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
                | |||||||||||  ||| |||||||| | ||  | ||||||||||||||||||||
Db        487 AATTATCATAAGTGGGTCAATATTTACTGTGGAGAAATAAATCCTTTTGTCTGCGAGGCA 546

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCCACCCCATCCC 420
              |||||||||||| ||||||  |||||||| ||||||||||||||||||||||||   ||
Db        547 TAGTCTGAAGATGCAGCTGTGTGAAGTCTGCAGAAGCAAGGAAGCCCCCCCACCCCCCACC 606

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTC 463
              ||   |||||      |   | ||  | | ||    |   |||||
Db        607 CCCCACCTGCTGCATCTGTAGCTGGGATCTGGTTCTGCTGCTC 649


RESULT 3
US-10-226-420-4
; Sequence 4, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 721
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)...(546)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk3
US-10-226-420-4
```

```
  Query Match             56.3%;  Score 306.2;  DB 13;  Length 721;
  Best Local Similarity   78.8%;  Pred. No. 5e-94;
  Matches  365;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   |||||||||| ||||||||||||   |||||  |||   ||||||||||  |||
Db        187 GATCAGCATTGCTACAGGGTCTTCAAACAACTCAAGACGTGGGACGATGCAGAGAGGTTC 246

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| ||  ||||||  | | ||  ||||||||  ||||||||||||||||||||  |||||||| |
Db        247 TGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTCTCTATCGAAAGCTCCGAAGAAGCAGCC 306

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||  ||||  ||  |   ||  ||||  |   |  |  |||||   ||  ||  ||||||||||
Db        307 TTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGGAGAGCCATTCTCTATATCTGGATCGGA 366

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              |||||||  ||||  ||||||  ||||||||||||     ||||||||||||||||||||  ||
Db        367 CTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGCGCGAAGTGGAGCGATGGCTCCAGCGTC 426

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||  ||||||||||  ||||||||||   |||||||| ||||    ||  |||
Db        427 AGTTATGAGAACTGGATTGAAGCAGAATCCAAAACATGTCTTGGGCTGCAACAAGGCACA 486

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
                | ||||||||||  |||  |||||||||  | ||  |  ||||||||||||||||||||
Db        487 AATTATCATAAGTGGGTCAATATTTACTGTGGAGAAATAAATCCTTTTGTCTGCGAGGCA 546

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCC 420
              ||||||||||||  ||||||  |||||||||  ||||||||||||||||||||||||    ||
Db        547 TAGTCTGAAGATGCAGCTGTGTGAAGTCTGCAGAAGCAAGGAAGCCCCCCACCCCCCACC 606

Qy        421 CCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTC 463
              ||   |||||      |  | ||  | | ||    |   |||||
Db        607 CCCCACCTGCTGCATCTGTAGCTGGGATCTGGTTCTGCTGCTC 649


RESULT 4
US-09-929-230-10
; Sequence 10, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (88)...(561)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk5
US-09-929-230-10

  Query Match             52.9%;  Score 288;  DB 10;  Length 725;
  Best Local Similarity   75.9%;  Pred. No. 8.8e-88;
  Matches  416;  Conservative    0;  Mismatches  110;  Indels   22;  Gaps    4;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   | |||||||||  |||| ||||||  |   |||||||||   |||||||| | |||
Db        190 GATCAGTATTGCTACAGGGTCATCAAACGACTCAAGACCTGGGACGATGCAGAGCGGTTC 249

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||  ||  ||||||  ||| ||  ||||||||||   |||  ||||||  |    | |||||||| |
Db        250 TGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCGTCTGTCGAAAACGATGAAGAAGCAGTC 309

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||  |||  |||||||  |  |||   |||  ||||||  |   |||||||||   | |||
Db        310 TTTCTGGCCCAGTTGGTCGCTGCGAACATAAAGCAAAACCAATACTATGTCTGGATTGGA 369

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||  ||||||||||| | ||||||||||||    ||||||||||||||||||||  ||
Db        370 CTGAGGATTCAAAACAAAGGACAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 429

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||    | ||  ||  | | ||||||||||||  |||||  ||  |  |  |||||||
Db        430 AGTTATGAGAACCTGGTTAAATCACATTCCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 489

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356
              | |||||  |  | |||  |  |||   |  ||||  |||  ||  ||  |  |  |||  ||||||  |
Db        490 GAGTTTCTTCAATGGTACAATACTGACTGCGAAGAAAAAAACCTTTTCGTCTGCAAGTTC 549

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414
                | ||  |||  |  |||||||  ||||   ||||||||||||||||||||||||  ||||||  |  |
Db        550 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAATCCCCCCCCAC 609

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473
              |  ||                |||||||||||||||  |||||  ||  |   ||  |  |||||||
Db        610 CCGCC-------------TGCCACAATCTCTGCTCTGCACTCTGTCATTCCATGGATGCT 656

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
              ||||||  ||||||||||||  ||||||||||||||||||  ||||   ||||||||||||
Db        657 CTCTGTGGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 716

Qy        532 CTAGCCTG 539
              |||||  ||
Db        717 CTAGCATG 724


RESULT 5
US-10-226-420-10
; Sequence 10, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (88)...(561)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk5
US-10-226-420-10

  Query Match             52.9%;  Score 288;  DB 13;  Length 725;
  Best Local Similarity   75.9%;  Pred. No. 8.8e-88;
  Matches  416;  Conservative    0;  Mismatches  110;  Indels   22;  Gaps    4;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   | |||||||||  |||| ||||||  |   |||||||||   |||||||| | |||
Db        190 GATCAGTATTGCTACAGGGTCATCAAACGACTCAAGACCTGGGACGATGCAGAGCGGTTC 249

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||  ||  ||||||  ||| ||  ||||||||||   |||  ||||||  |    | |||||||| |
Db        250 TGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCGTCTGTCGAAAACGATGAAGAAGCAGTC 309

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
```

```
Db        310 TTTCTGGCCCAGTTGGTCGCTGCGAACATAAAGCAAAACCAATACTATGTCTGGATTGGA 369

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        370 CTGAGGATTCAAAACAAAGGACAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 429

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        430 AGTTATGAGAACCTGGTTAAATCACATTCCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 489

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356

Db        490 GAGTTTCTTCAATGGTACAATACTGACTGCGAAGAAAAAAACCTTTTCGTCTGCAAGTTC 549

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414

Db        550 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAATCCCCCCCCAC 609

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCT-TTGCTCAACGGATGCT 473

Db        610 CCGCC------------TGCCACAATCTCTGCTCTGCACTCTGTCATTCCATGGATGCT 656

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531

Db        657 CTCTGTGGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGC 716

Qy        532 CTAGCCTG 539

Db        717 CTAGCATG 724
```

RESULT 6
US-09-929-230-7
; Sequence 7, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(434)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk4
US-09-929-230-7

Query Match 52.5%; Score 285.8; DB 10; Length 580;
Best Local Similarity 75.2%; Pred. No. 4.4e-87;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         63 GATCAGTATTGCTACAGGGTCATCAAACAACTCAGGACGTGGGACGATGCAGAGAGGTTC 122

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        123 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTCTCTATTGAAAGCGACGGAGAAGCAGCC 182

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        183 TTTGTGGCCCAGCTGGTCGCTGAGAACATCAAGCAAAACAAATATGATGTCTGGATCGGA 242

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        243 CTGAGGATTCAAGGCGAAGAGAAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 302

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        303 AATTATGAGAACCTGATTAAACATGCGACCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 362

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356

Db        363 GGGTTTCGCACGTGGCGCAATGTTCACTGTACACAACAAAATCTTTTCATGTGCAAGTTC 422

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414

Db        423 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 482

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474

Db        483 CACCGC------------------------------CACCCTTTGCTCAACGGATGCTC 511

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532

Db        512 TCTGTAGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAACTCTTCC 571

Qy        533 TAGCCTGAA 541

Db        572 TAGCCTGGA 580
```

RESULT 7
US-10-226-420-7
; Sequence 7, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 580
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(434)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk4
US-10-226-420-7

Query Match 52.5%; Score 285.8; DB 13; Length 580;
Best Local Similarity 75.2%; Pred. No. 4.4e-87;
Matches 413; Conservative 0; Mismatches 97; Indels 39; Gaps 3;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         63 GATCAGTATTGCTACAGGGTCATCAAACAACTCAGGACGTGGGACGATGCAGAGAGGTTC 122

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        123 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTCTCTATTGAAAGCGACGGAGAAGCAGCC 182

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        183 TTTGTGGCCCAGCTGGTCGCTGAGAACATCAAGCAAAACAAATATGATGTCTGGATCGGA 242

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
```

```
Db        243 CTGAGGATTCAAGGCGAAGAGAAGCAATGCAGCACGAAGTGGAGCGATGGCTCCAGCGTC 302

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        303 AATTATGAGAACCTGATTAAACATGCGACCAAAAAGTGTTTTGGGCTGAAAAAAGAGACA 362

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356

Db        363 GGGTTTCGCACGTGGCGCAATGTTCACTGTACACAACAAAATCTTTTCATGTGCAAGTTC 422

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414

Db        423 CCGCCAGAGTGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 482

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474

Db        483 CACCGC------------------------------CACCCTTTGCTCAACGGATGCTC 511

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532

Db        512 TCTGTAGCTGGATCTGGTTCTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAACTCTTCC 571

Qy        533 TAGCCTGAA 541

Db        572 TAGCCTGGA 580
```

RESULT 8
US-09-969-763-2
; Sequence 2, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
; APPLICANT: FUKUCHI, NAOYUKI
; APPLICANT: KITO, MORIKAZU
; APPLICANT: KAYAHARA, TAKASHI
; APPLICANT: FUTAKI, FUMIE
; APPLICANT: ISHIKAWA, KOHKI
; APPLICANT: SUZUKI, EIICHIRO
; APPLICANT: GONDOH, KEIKO
; APPLICANT: SHIMBA, NOBUHISA
; APPLICANT: YAMADA, NAOYUKI
; TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
; FILE REFERENCE: 214760US0
; CURRENT APPLICATION NUMBER: US/09/969,763
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: JP 2000-305279
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Crotalus harridus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (66)..(512)
; OTHER INFORMATION:
US-09-969-763-2

Query Match 41.6%; Score 226.4; DB 10; Length 690;
Best Local Similarity 70.5%; Pred. No. 1.1e-66;
Matches 389; Conservative 0; Mismatches 126; Indels 37; Gaps 5;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db        168 GATCGGTATTGCTACAAGCCCTTCAAACAAGAGATGACCTGGGCCGATGCAGAGAGGTTC 227

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        228 TGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTCTCTGTCGAAACCGCCCTAGAAGCATCC 287

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        288 TTTGTGGACAATGTGCTCTATGCGAACAAAGAGTACCTCACACGTTATATCTGGATTGGA 347

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        348 CTGAGGGTTCAAAACAAAGGACAGCCAT--------------------GCTCCAGCATC 386

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        387 AGTTATGAGAACCTGGTT------GACCCATTTGAATGTTTTATGGTGAGCAGAGACACA 440

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGA---- 356

Db        441 AGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAACAACAACATTCTTTCATATGCAAGTTC 500

Qy        357 --GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 414

Db        501 ACGCGACCACGTTAAGATCCGGCTGTGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCT 560

Qy        415 CATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTC 474

Db        561 C--TCCCCACCCCCCACCTTCCGCAATCTCTGCTCTTCCCCCTTTGCTCAGTGGATGCTC 618

Qy        475 TCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCC 532

Db        619 TCTGTAGCCGGATCTGGGTTTTCTGCTCCAGATGGGTCAGAAGATCCAATAAATTCTGCC 678

Qy        533 TAGCCTGAAAAA 544

Db        679 TACCCAAAAAAA 690
```

RESULT 9
US-09-929-230-6
; Sequence 6, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
; NAME/KEY: misc_feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
; LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
; LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
; LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-6

Query Match 34.3%; Score 186.6; DB 10; Length 456;
Best Local Similarity 47.1%; Pred. No. 4e-53;
Matches 169; Conservative 86; Mismatches 104; Indels 0; Gaps 0;

```
Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         97 GAYCARCAYTGYTAYMGNGTNTTYAARCARYTNAARACNTGGGAYGAYGCNGARMGNTTY 156

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
```

```
Db        157 TGYWSNGARCARGCNGARGGNGGNCAYYTNGTNWSNATHGARWSNWSNGARGARGCNGCN 216

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        217 TTYGTNGCNCARYTNGTNCCNGARAAYMGNMGNMGNGCNATHYTNTAYATHTGGATHGGN 276

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        277 YTNMGNGTNCARGGNAARGARAARCARTGYWSNGCNAARTGGWSNGAYGGNWSNWSNGTN 336

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        337 WSNTAYGARAAYTGGATHGARGCNGARWSNAARACNTGYYTNGGNYTNCARCARGGNACN 396

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGC 359

Db        397 AAYTAYCAYAARTGGGTNAAYATHTAYTGYGGNGARATHAAYCCNTTYGTNTGYGARGC 455


RESULT 10
US-10-226-420-6
; Sequence 6, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
;   LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
;   LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
;   LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-6

  Query Match             34.3%;  Score 186.6;  DB 13;  Length 456;
  Best Local Similarity   47.1%;  Pred. No. 4e-53;
  Matches  169;  Conservative   86;  Mismatches  104;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         97 GAYCARCAYTGYTAYMGNGTNTTYAARCARYTNAARACNTGGGAYGAYGCNGARMGNTTY 156

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        157 TGYWSNGARCARGCNGARGGNGGNCAYYTNGTNWSNATHGARWSNWSNGARGARGCNGCN 216

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        217 TTYGTNGCNCARYTNGTNCCNGARAAYMGNMGNMGNGCNATHYTNTAYATHTGGATHGGN 276

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        277 YTNMGNGTNCARGGNAARGARAARCARTGYWSNGCNAARTGGWSNGAYGGNWSNWSNGTN 336

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        337 WSNTAYGARAAYTGGATHGARGCNGARWSNAARACNTGYYTNGGNYTNCARCARGGNACN 396

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGC 359

Db        397 AAYTAYCAYAARTGGGTNAAYATHTAYTGYGGNGARATHAAYCCNTTYGTNTGYGARGC 455


RESULT 11
US-09-929-230-9
; Sequence 9, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;   NAME/KEY: misc_feature
;   LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;   LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;   LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;   LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9

  Query Match             32.2%;  Score 175;  DB 10;  Length 432;
  Best Local Similarity   45.7%;  Pred. No. 3.7e-49;
  Matches  163;  Conservative   85;  Mismatches  109;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Db         61 GAYCARTAYTGYTAYMGNGTNATHAARCARYTNMGNACNTGGGAYGAYGCNGARMGNTTY 120

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Db        121 TGYWSNGARCARGCNAARGGNGGNCAYYTNGTNWSNATHGARWSNGAYGGNGARGCNGCN 180

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Db        181 TTYGTNGCNCARYTNGTNGCNGARAAYATHAARCARAAYAARTAYGAYGTNTGGATHGGN 240

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        241 YTNMGNATHCARGGNGARGARAARCARTGYWSNACNAARTGGWSNGAYGGNWSNWSNGTN 300

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Db        301 AAYTAYGARAAYYTNATHAARCAYGCNACNAARAARTGYTTYGGNYTNAARAARGARACN 360

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG 357

Db        361 GGNTTYMGNACNTGGMGNAAYGTNCAYTGYACNCARCARAAYYTNTTYATGTGYAAR 417


RESULT 12
US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;   LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;   LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;   LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-9

  Query Match             32.2%;  Score 175;  DB 13;  Length 432;
  Best Local Similarity   45.7%;  Pred. No. 3.7e-49;
  Matches  163;  Conservative   85;  Mismatches  109;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||    : |:||:||::  ||  |:||:||::  :  || |||   ||:|| ||::| ||:
Db         61 GAYCARTAYTGYTAYMGNGTNATHAARCARYTNMGNACNTGGGAYGAYGCNGARMGNTTY 120

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||:::  |:||:|  || || || ||::| || :: ||:||:::   :|| ||:|| |
Db        121 TGYWSNGARCARGCNAARGGNGGNGCAYYTNGTNWSNATHGARWSNGAYGGNGARGCNGCN 180

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||:|| |  ||::|  | ||  |:|| ||:||:  :  :||:  : |:|| ||||:||
Db        181 TTYGTNGCNCARYTNGTNGCNGARAAYATHAARCARAAYAARTAYGAYGTNTGGATHGGN 240

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              :| :|   :||:    |:||:||:||:||::: |   |:|||:: ||:|| :: ::  |
Db        241 YTNMGNATHCARGGNGARGARAARCARTGYWSNACNAARTGGWSNGAYGGNWSNWSNGTN 300

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              | :||:||:||::  ||: |: | |   | ||:||:||: |:||  |  | | :||:||
Db        301 AAYTAYGARAAYYTNATHAARCAYGCNACNAARAARTGYTTYGGNYTNAARAARGARACN 360

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG 357
              || ||::  |  |||   ||: |  |:||:   ||:||: |::  ||: | ||: |:
Db        361 GGNTTYMGNACNTGGMGNAAYGTNCAYTGYACNCARCARAAYYTNTTYATGTGYAAR 417


RESULT 13
US-09-929-230-12
; Sequence 12, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
;   LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
;   LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
;   LOCATION: 402, 411, 426, 447, 453, 465, 468
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-12

  Query Match             30.5%;  Score 166;  DB 10;  Length 474;
  Best Local Similarity   43.4%;  Pred. No. 4.8e-46;
  Matches  155;  Conservative   88;  Mismatches  114;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||    : |:||:||::  ||  |:||::  :  ||:|| |||   ||:|| ||::| ||:
Db        103 GAYCARTAYTGYTAYMGNGTNATHAARMGNYTNAARACNTGGGAYGAYGCNGARMGNTTY 162

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||:::  |:||:|  || || || ||::| |  ::  | ||:| :  :| :||:|| |
Db        163 TGYWSNGARCARGCNAARGGNGGNGCAYYTNGCNWSNGTNGARAAYGAYGARGARGCNGTN 222

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||: | |  ||::|  | ||    || ||:||:  :  : |:  : |:|| |||||:||
Db        223 TTYYTNGCNCARYTNGTNGCNGCNAAAYATHAARCARAAYCARTAYTAYGTNTGGATHGGN 282

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              :| :|   :||:||:||:|  |:|[|:||::: |   |:|||:: ||:|| :: ::  |
Db        283 YTNMGNATHCARAAYAARGGNCARCARTGYWSNACNAARTGGWSNGAYGGNWSNWSNGTN 342

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              :: ||:||:||::  |  |:  | :: ||:||:||: |:||  |  | | :||:||
Db        343 WSNTAYGARAAYYTNGTNAARWSNCAYWSNAARAARTGYTTYGGNYTNAARAARGARACN 402

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG 357
              | :||::   |:||| | ||:    |:||:||: |: |: |::  ||:|| ||: |:
Db        403 GARTTYYTNCARTGGTAYAAYACNGAYTGYGARGARAARAAYYTNTTYGTNTGYAAR 459


RESULT 14
US-10-226-420-12
; Sequence 12, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
;   LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
;   LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
;   OTHER INFORMATION: n = A,T,C or G
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
;   LOCATION: 402, 411, 426, 447, 453, 465, 468
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-12

  Query Match             30.5%;  Score 166;  DB 13;  Length 474;
  Best Local Similarity   43.4%;  Pred. No. 4.8e-46;
  Matches  155;  Conservative   88;  Mismatches  114;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||    : |:||:||::  ||  |:||::  :  ||:|| |||   ||:|| ||::| ||:
Db        103 GAYCARTAYTGYTAYMGNGTNATHAARMGNYTNAARACNTGGGAYGAYGCNGARMGNTTY 162

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||:::  |:||:|  || || || ||::| |  ::  | ||:| :  :| :||:|| |
Db        163 TGYWSNGARCARGCNAARGGNGGNCAYYTNGCNWSNGTNGARAAYGAYGARGARGCNGTN 222

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||: | |  ||::|  | ||    || ||:||:  :  : |:  : |:|| |||||:||
Db        223 TTYYTNGCNCARYTNGTNGCNGCNAAYATHAARCARAAYCARTAYTAYGTNTGGATHGGN 282

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              :| :|   :||:||:||:|   |:||:||::: |   |:|||:: ||:|| :: ::  |
Db        283 YTNMGNATHCARAAYAARGGNCARCARTGYWSNACNAARTGGWSNGAYGGNWSNWSNGTN 342

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              :: ||:||:||::   | |:    | :: ||:||:||: |:||  |  | | :||:||
Db        343 WSNTAYGARAAYYTNGTNAARWSNCAYWSNAARAARTGYTTYGGNYTNAARAARGARACN 402

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG 357
              | :||::   |:||| | ||:    |:||:||: |: |: |::  ||:|| ||: |:
Db        403 GARTTYYTNCARTGGTAYAAYACNGAYTGYGARGARAARAAYYTNTTYGTNTGYAAR 459


RESULT 15
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(455)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk2
US-09-929-230-1

  Query Match             15.1%;  Score 82.4;  DB 10;  Length 483;
  Best Local Similarity   53.9%;  Pred. No. 2.3e-17;
  Matches  220;  Conservative    0;  Mismatches  176;  Indels   12;  Gaps    2;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||   ||||||||||||||||||||   ||   || ||||||   |||||||||| |||
Db         87 GATCAGCATTGCTACAAGGTCTTCAGTGAACTCAAAACCTGGGATGATGCAGAGAGTTTC 146

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              |||   |  |||   |  |   | |  |||| ||| ||| | |||   | ||||| ||
Db        147 TGCTACACACAGCACAGAGACAGCCGCCTGGCCTCCATCCACAGCAGTGAAGAAGAAGCT 206

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              |||||||| |  ||    | || |   | || |   |      | |||||||||
Db        207 TTTGTGGGCAAACTGGCCTCCCAAACTTTGAAATTCACTTCCA------TGTGGATCGGA 260

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||  | || |   ||||      ||||||      |||||||||||| | |||   |
Db        261 CTGAAAGATCTATGGAAAG------AATGCAAATGGCAGTGGAGCGATGACACCAAACTG 314

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||  |     |||| |  |  |   |      |      | |  ||  | | ||| |
Db        315 GACTACAAAGCCTGGACTCGAAGACCCTATTGTACAGTAATGGTAGTCAAGACAGATAGG 374

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
              |||              ||   |      |    |   |    | | |||||
Db        375 ATCTTTTGGTTCAATAGAGGTTGCGAAAAGACTGTATCTTTTGTCTGCAAGTTCCAGGCA 434

Qy        361 TAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCC 408
              |||||  |||||| |||| |||||||| |||||||   |  |||
Db        435 CGGTCTGGAGATCCGGCTGTGTGAAGTCTGCAGAAGCAAAAGAAGCCC 482


Search completed: December  8, 2003, 13:54:34
Job time : 348 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 10:20:33 ; Search time 2261 Seconds
(without alignments)
5847.694 Million cell updates/sec

Title: US-09-938-114-1
Perfect score: 544
Sequence: 1 GAAGGGCATTGCTACAAGGT.....ATTCTGCCTAGCCTGAAAAA 544

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	302	55.5	461	12	BM401683	BM401683 PH012R Sn
	2	233.8	43.0	418	12	BM401662	BM401662 JL2G06F S
	3	222.4	40.9	556	12	BM401668	BM401668 JL2H08F S
	4	205	37.7	618	12	BM401682	BM401682 PH012F Sn
	5	203.8	37.5	501	12	BM401648	BM401648 JL2B03F S
	6	195.6	36.0	370	12	BM401634	BM401634 JL1G02F S
	7	186.8	34.3	458	12	BM401641	BM401641 JL1H06F S
	8	186	34.2	460	12	BM401460	BM401460 JH1D06F S
	9	180.6	33.2	511	12	BM401631	BM401631 JL1F11F S
c	10	172.8	31.8	657	12	BM401418	BM401418 GH061R Sn
	11	161.2	29.6	406	12	BM401598	BM401598 JH4F11F S
	12	159.6	29.3	410	12	BM401605	BM401605 JH4H07F S
	13	156.2	28.7	453	12	BM401653	BM401653 JL2D01F S
	14	152.2	28.0	440	12	BM401459	BM401459 JH1D05F S
	15	151.2	27.8	456	12	BM401659	BM401659 JL2F01F S
	16	149.4	27.5	440	12	BM401637	BM401637 JL1G11F S
	17	146.8	27.0	654	12	BM401401	BM401401 GH025F Sn
c	18	134	24.6	341	12	BM401402	BM401402 GH025RB S
	19	129.4	23.8	248	12	BM401663	BM401663 JL2G08F S
	20	120.6	22.2	405	12	BM401607	BM401607 JH4H10F S
	21	112.4	20.7	574	12	BM401417	BM401417 GH061F Sn
	22	111.2	20.4	388	12	BM401625	BM401625 JL1E05F S
	23	99.8	18.3	401	12	BM401610	BM401610 JL1A11F S
	24	99.8	18.3	633	12	BM401686	BM401686 PH015FB S
	25	99.6	18.3	365	12	BM401665	BM401665 JL2H01F S
	26	82	15.1	510	12	BM401667	BM401667 JL2H07F S
	27	77.4	14.2	321	12	BM401611	BM401611 JL1B01F S
	28	77.4	14.2	378	12	BM401620	BM401620 JL1C11F S
	29	75	13.8	475	12	BM401621	BM401621 JL1D02F S
	30	75	13.8	504	12	BM401628	BM401628 JL1E12F S
	31	75	13.8	512	12	BM401495	BM401495 JH2B09F S
	32	74	13.6	301	12	BM401654	BM401654 JL2D07F S
	33	74	13.6	378	12	BM401618	BM401618 JL1C07F S
	34	71	13.1	391	12	BM401469	BM401469 JH1E09F S
	35	70.4	12.9	315	12	BM401503	BM401503 JH2D02F S
	36	67.6	12.4	497	12	BM401552	BM401552 JH3E12F S
	37	63.4	11.7	374	12	BM401520	BM401520 JH2G02F S
	38	63	11.6	383	12	BM401463	BM401463 JH1D11F S
	39	63	11.6	406	12	BM401586	BM401586 JH4D11F S
	40	63	11.6	414	12	BM401532	BM401532 JH3A10F S
	41	63	11.6	501	12	BM401527	BM401527 JH2H09F S
	42	62.6	11.5	407	12	BM401658	BM401658 JL2E09F S
	43	60.4	11.1	570	12	BM401652	BM401652 JL2C09F S
	44	58.6	10.8	343	12	BM401656	BM401656 JL2E04F S
	45	57	10.5	307	12	BM401567	BM401567 JH4A07F S

ALIGNMENTS

RESULT 1
BM401683/c
LOCUS BM401683 461 bp mRNA linear EST 01-MAY-2002
DEFINITION PH012R Snake Bothrops insularis library IL2 Bothrops insularis cDNA 3' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401683
VERSION BM401683.1 GI:20376311
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 461)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505

```
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL09A (see Reference)
            Seq primer: M13R.
FEATURES              Location/Qualifiers
     source           1. .461
                      /organism="Bothrops insularis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8723"
                      /tissue_type="venom glands"
                      /clone_lib="Snake Bothrops insularis library IL2"
                      /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                      RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                      venom glands were primed with oligo-(dT) and reverse
                      transcribed to cDNA using Superscript Plasmid System for
                      cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                      were selected by size (350-600 pb and up 600 pb) in
                      agarose gel electrophoresis, linked to Eco RI adapters and
                      directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                      were generated from random clones and grouped in unique
                      sequences. The putative identification of each EST or
                      cluster was obtained through Blast searches (e-value <
                      e-05)."
BASE COUNT      104 a    125 c    107 g    120 t      5 others
ORIGIN

  Query Match             55.5%;  Score 302;  DB 12;  Length 461;
  Best Local Similarity   82.2%;  Pred. No. 3.2e-71;
  Matches  382;  Conservative    0;  Mismatches   75;  Indels    8;  Gaps    3;

Qy         74 TGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGT 133
              |||| || || |||||||| |||||||||||| |||||||||||| |||||||| ||||
Db        461 TGAATGGCGGCCATCTGGTNTCTATCGAAAGCGCCGGAGAAGCAGCCTTTGTGGCCCAGC 402

Qy        134 TGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAA 193
              || | | |||| | | |       | || |  |||||||||||||| |||||| ||||
Db        401 TGGTCGNTGAGAACAAACACAGTCGCGAATTGAATGTCTGGATCGGAATGAGGGTTCAAG 342

Qy        194 ACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATT 253
              ||||||||||||||||||      |||||||||||||||||| |||||||||||| |
Db        341 GCAAAGAAAAGCAATGCAGC---TCGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACT 285

Qy        254 GGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGT 313
              ||||| ||| ||||| |||| ||| ||||| || | | |||  |||||||| |||||
Db        284 GGATTAAAGCAGAATTAAAAACATGTATTGGGTTGAAAAAAGATGCAGGGTTTCGTAAGT 225

Qy        314 GGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCATAGTCTGAAGATC 373
              ||   ||| || |||||||||||| |  |||||| |||||||||| ||||||||||||
Db        224 GGAGCAATNTTAACTGTGAACAACGAATTCCTTTNGTCTGCGAGGTATAGTCTGAAGATG 165

Qy        374 CAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCCCATCCCCCAACCCTGCCTA 433
              ||||||| |||||||||||||||||||||||||||||||||   | ||| ||||   |
Db        164 CAGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC---CTCCCCACCCCCGCCT 108

Qy        434 GCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTCTGTAGCTGGATCTGGTG 493
              ||| |||||||||| ||| ||||| ||| |||||||||||||||||||||||||||||
Db        107 GCCGCAATCTCTGCTCTGCCCCCTTCGCTGAACGGATGCTCTCTGTAGCTGGATCTGGTT 48

Qy        494 TTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTAGC 536
              |||||||||||||||||| ||||   |||||||| ||||||||
Db         47 TTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTNTGCCTAGC 3


RESULT 2
BM401662
LOCUS       BM401662                 418 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2G06F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401662
VERSION     BM401662.1  GI:20376290
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES              Location/Qualifiers
     source           1. .418
                      /organism="Bothrops insularis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8723"
                      /tissue_type="venom glands"
                      /clone_lib="Snake Bothrops insularis library IL3"
                      /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                      RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                      venom glands were primed with oligo-(dT) and reverse
                      transcribed to cDNA using Superscript Plasmid System for
                      cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                      were selected by size (350-600 pb and up 600 pb) in
                      agarose gel electrophoresis, linked to Eco RI adapters and
                      directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                      were generated from random clones and grouped in unique
                      sequences. The putative identification of each EST or
                      cluster was obtained through Blast searches (e-value <
                      e-05)."
BASE COUNT      116 a     90 c    122 g     90 t
ORIGIN

  Query Match             43.0%;  Score 233.8;  DB 12;  Length 418;
  Best Local Similarity   75.7%;  Pred. No. 1.1e-52;
  Matches  305;  Conservative    0;  Mismatches   92;  Indels    6;  Gaps    1;

Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
              || ||||   |||||||||| |||||| || |||||| ||| || || ||||||||||
Db         13 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 72

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
              ||||||||||   | |||||||||||||||| |||| || | |||| ||| |||  |
Db         73 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 132

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
              | |||||  | ||||||||||||||||||  | ||||||||  || ||||||||||
Db        133 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 192

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275
              ||||||||||  |||||||  |||||||||||||||   | ||  || | | |  |||||
Db        193 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 252

Qy        276 GTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACA 335
              |||| ||| | || | | | ||  |||||||| ||||||||  ||| || |||||| | |
Db        253 GTGTTTGCGCTGGAAAAAAAGCAAGGGTTTCGTAAGTGGGTCAATATTGACTGTGTAGA 312

Qy        336 ACAAGATCCTTTTGTCTGCGAG------GCATAGTCTGAAGATCCAGCTGATTGAAGTCT 389
              |  | ||||||| || ||| ||         |  |  | ||||||| ||||| ||||||||
Db        313 AGGAAATCCTTTCGTGTGCAAGTTCATACGACCGCGTTAAGATCCGGCTGAGTGAAGTCT 372

Qy        390 GGAGAAGCAAGGAAGCCCCCCACCCCATCCCCCAACCCTGCCT 432
              |||||||||||| ||||||||||||      || |||| ||||
Db        373 GGAGAAGCAAGGGAGCCCCCCACCCCCCTTACCCACCCCGCCT 415
```

```
RESULT 3
BM401668
LOCUS       BM401668                556 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2H08F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401668
VERSION     BM401668.1  GI:20376296
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      146 a    126 c    154 g    129 t      1 others
ORIGIN

  Query Match             40.9%;  Score 222.4;  DB 12;  Length 556;
  Best Local Similarity   74.3%;  Pred. No. 1.5e-49;
  Matches  295;  Conservative    0;  Mismatches   97;  Indels    5;  Gaps    1;

Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
              ||  ||||   |||||||||||| |||||| || |||||| ||| || || ||||||||||
Db        159 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 218

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
              |||||||||   | ||||||||||||||||| |||| || | |||| ||| |||  |
Db        219 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 278

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
              | |||||  | |||||||||||||||||||  | |||||||| || ||||||||||
Db        279 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 338

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275
              ||||||||||  ||||||| ||||||||||||  | ||  || | |    |||||
Db        339 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 398

Qy        276 GTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACA 335
              |||| ||| | || | | | ||  |||||||| |||||||  ||| || |||||| | |
Db        399 GTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCGTAAGTGGGTCAATATTGACTGTGTAGA 458

Qy        336 ACAAGATCCTTTTGTCTGCGAG-----GCATAGTCTGAAGATCCAGCTGATTGAAGTCTG 390
              | | |||||||| || ||| ||       |    |   |    | ||||| ||||||||||
Db        459 AGGAAATCCTTTCGTGTGCAAGTTCATACGACCGCGTTAAGATCCGCTGAGTGAAGTCTG 518

Qy        391 GAGAAGCAAGGAAGCCCCCCACCCCATCCCCCAACCC 427
              ||||| |||||||||||||| |||| |  |||  |||
Db        519 GAGAANCAAGGAAGCCCCCCCCCCCCTTACCCCCCCC 555


RESULT 4
BM401682
LOCUS       BM401682                618 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  PH012F Snake Bothrops insularis library IL2 Bothrops insularis cDNA
            5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401682
VERSION     BM401682.1  GI:20376310
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL09A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL2"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      157 a    124 c    180 g    149 t      8 others
ORIGIN

  Query Match             37.7%;  Score 205;  DB 12;  Length 618;
  Best Local Similarity   77.1%;  Pred. No. 8.3e-45;
  Matches  297;  Conservative    0;  Mismatches   81;  Indels    7;  Gaps    4;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||| || ||||| | |||||||| || | || | ||||   |||||| |||| |||
Db        241 GAAGGGCGTTCCTACAGGTTCTTCAAAGAAGCGAAAAACTGGGACGATGCANAGAGGTTC 300
```

```
Qy          61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
               ||| || || ||||||| || || ||||||||||||||||||||| ||||||||||| |
Db         301 TGCTCGGAGGAGGTGAATGGCGGCCATCTGGTCTCTATCGAAAGCGCCGGAGAAGCAGCC 360

Qy         121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
               ||||||| |||| || | ||| |||| | | |       | || |  ||||||||||||||
Db         361 TTTGTGGCCCAGCTGGTCGCTGAGAACAAACACAGTCGCGAATTGAATGTCTGGATCGGA 420

Qy         181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
               |||||| |||| |||||||||  |     |||    |||||||||||||||| | ||
Db         421 ATGAGGGTTCAAGGCAAAGAAAACAA----TGCANCTCGTGGAGCGATGGCTCCA-CGTC 475

Qy         241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
               |||||||||| |||||| ||| | |||  |||| |||  ||||| || | | |||  ||
Db         476 AGTTATGAGAACTGGATTAAAGCANAATTAAAAACATGTATTGGGTTGAAAAAAGATGCA 535

Qy         301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360
               ||||||| |||||||   ||| || |||||||||||| | || | ||| ||||||||| |
Db         536 GGGTTTCNTAAGTGGAGCAATCTTAACTGTGAACAACNA-ATTCCTTTCTCTGCGAGGTA 594

Qy         361 TAGTCTGAAGATCCAGCTGATTGAA 385
               ||  |||||||| |||||| ||||
Db         595 TA-NCTGAAGATGCAGCTGAGTGAA 618
```

```
RESULT 5
BM401648
LOCUS       BM401648                501 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2B03F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401648
VERSION     BM401648.1  GI:20376276
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      134 a    102 c    139 g    122 t      4 others
ORIGIN
```

```
  Query Match           37.5%;  Score 203.8;  DB 12;  Length 501;
  Best Local Similarity   75.8%;  Pred. No. 1.7e-44;
  Matches  250;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy          36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
               || ||||   |||||||||||| |||||| || |||||| ||| || || ||||||||||||
Db         161 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 220

Qy          96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
               ||||||||||    | ||||||||||||||||| |||| || | |||| ||| |||  |
Db         221 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 280

Qy         156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
               | |||||  | |||||||||||||||||||  | |||||||| || ||||||||||
Db         281 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 340

Qy         216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275
               ||||||||||  |||||||  |||||||||||||   | ||  || | |    |||||
Db         341 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 400

Qy         276 GTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAACA 335
               |||| ||| | || | | | ||  |||||||| ||||||||  ||| || ||||||| | |
Db         401 GTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCNTAAGTGGGTCAATATTGACTGTGTANA 460

Qy         336 ACAAGATCCTTTTGTCTGCGAGGCATAGTC 365
               |  | |||||||  | ||| |  ||| | |
Db         461 AGGAAATCCTTTCNTGTGCAATTCATCGAC 490
```

```
RESULT 6
BM401634
LOCUS       BM401634                370 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL1G02F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401634
VERSION     BM401634.1  GI:20376262
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
```

transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 96 a 92 c 92 g 85 t 5 others
ORIGIN

Query Match 36.0%; Score 195.6; DB 12; Length 370;
Best Local Similarity 76.2%; Pred. No. 2.6e-42;
Matches 279; Conservative 0; Mismatches 78; Indels 9; Gaps 3;

```
Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||  | ||||||||  || ||||||||||    ||||||||||  |||||||  ||
Db          1 CTGAGGATTGAAAACAAAAAACAGCAATGCAGCTNGAAGTGGAGCGATTACTCCAGCGTC 60

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300
              ||||||||||    | ||  || | |    |||||||| ||| | || | | | ||  |
Db         61 AGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAAGTGTTTTGCGCTGGAAAAAAAGCAA 120

Qy        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAG--- 357
              |||||||  ||||||||  ||| || |||||| | ||  | |||||||  || ||| ||
Db        121 GGGTTTCGTAAGTGGGTCAATATTGACTGTGTAGAAGGAAATCCTTTCGTGTGCAAGTTC 180

Qy        358 ---GCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCA-CC 413
                 |  |  | |||||||  |||||  ||||||||||||||||||||||||||||||  ||
Db        181 ATACGACCGCGTTAAGATCCGGCTGAGTGAAGTCTGGAGAAGCAAGGAAGCCCCCCACCC 240

Qy        414 CCATCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCT 473
              ||  |  |||  ||||    |  ||  |||  |||||||  |||  |||||  |||  ||||||||||
Db        241 CCCTTACCCCACCCCCGCCTGCTGCAACCTCTGCTCTGCCCCCTTCGCTGAACGGATGCT 300

Qy        474 CTCTGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGC 531
              |||||||  ||||||||||  ||||||||||||||||||||  |||   ||||  ||||||||
Db        301 CTCTGTANCTGGATCTGGTTTTGCTGCTCCTGATGGGCCANAAGGTCCAATNAATTCTGC 360

Qy        532 CTAGCC 537
              |||  ||
Db        361 CTAACC 366
```

RESULT 7
BM401641
LOCUS BM401641 458 bp mRNA linear EST 01-MAY-2002
DEFINITION JL1H06F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401641
VERSION BM401641.1 GI:20376269
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 458)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .458
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 122 a 90 c 130 g 111 t 5 others
ORIGIN

Query Match 34.3%; Score 186.8; DB 12; Length 458;
Best Local Similarity 76.2%; Pred. No. 6.7e-40;
Matches 227; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
              ||  ||||   |||||| | || |||||| || |||||| ||| || || ||||||||||
Db        161 GAGCTGGTATGATGCANANAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 220

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
              |||||||||   | ||||||||||||||||| |||| || | |||| ||| |||  |
Db        221 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 280

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
              | |||||  | |||||||||||||||||||  | ||||||| || ||||||||||
Db        281 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 340

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275
              ||||||||||  ||||||| |||||||||||||  | ||  || | |   |||||
Db        341 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 400

Qy        276 GTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAA 333
              |||| ||| | || | | | ||  | |||||| ||||||||  ||| || |||||| |
Db        401 GTGTTTTGCGCTGGAAAAAAAGCAANGGTTTCNTAAGTGGGTCAATATTGACTGTGTA 458
```

RESULT 8
BM401460
LOCUS BM401460 460 bp mRNA linear EST 01-MAY-2002
DEFINITION JH1D06F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401460
VERSION BM401460.1 GI:20376088
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 460)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan

```
          Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
          Tel: 55 11 37 26 7222 ext. 2083
          Fax: 55 11 37 26 1505
          Email: hoplee@usp.br
          This EST corresponds to cluster BITL13A (see Reference)
          Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .460
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      106 a    106 c    145 g    103 t
ORIGIN

  Query Match             34.2%;  Score 186;  DB 12;  Length 460;
  Best Local Similarity   77.6%;  Pred. No. 1.1e-39;
  Matches  225;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              |||||| |||||||| ||||||||   || |  | | |||| | ||||||||||| |||
Db        170 GAAGGGAGTTGCTACAGGGTCTTCACTGAACCGCAAAACTGGGCCGATGCAGAGAGGTTC 229

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || |||||| ||| || || |||||||||||||||||||||    ||||||||||
Db        230 TGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTCTATCGAAAGCGAAGGAGAAGCAGAC 289

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
              ||||||| |||| || | ||| ||| |||| |   | || ||  ||||||||||||||||
Db        290 TTTGTGGCCCAGCTGGTCGCTGAGAGGATAGACAGACCCCAATCCCATGTCTGGATCGGA 349

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240
              ||||||| | || |||||| | ||||||||||     | ||||||||||||||| || ||
Db        350 CTGAGGGGTGAACACAAAGGACAGCAATGCAGCTCGAACTGGAGCGATGGCTCCTGCGTC 409

Qy        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAAGTGTCTTGGGGTGCA 290
              |||||||||   | |||||  |  | | |||||  ||| ||| | || |
Db        410 TGTTATGAGAACCTGGTTGAACGAAGAACGAAAAAATGTTTTGTGTTGGA 459


RESULT 9
BM401631
LOCUS       BM401631                511 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL1F11F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401631
VERSION     BM401631.1  GI:20376259
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      135 a    105 c    149 g    119 t      3 others
ORIGIN

  Query Match             33.2%;  Score 180.6;  DB 12;  Length 511;
  Best Local Similarity   76.6%;  Pred. No. 3.4e-38;
  Matches  219;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
              || ||||   |||||||||||| |||||| || |||||| ||| || || ||||||||||||
Db        225 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 284

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
              ||||||||||   | ||||||||||||||||| |||  || | |||| ||| |||   |
Db        285 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCANCTGGTCGCTCCGAACATAGGGAA 344

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
              | |||||  | |||||||||||||||||||  | |||||||| || | |||||||||
Db        345 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACANCAATGCAGCTC 404

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275
              ||||||||||  |||||||  |||||||||||||||   | ||  || | |    |||||
Db        405 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 464

Qy        276 GTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAAT 321
              |||| ||| | || | | | ||  |||||||| ||||||||  |||
Db        465 GTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCNTAAGTGGGTCAAT 510


RESULT 10
BM401418/c
LOCUS       BM401418                657 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  GH061R Snake Bothrops insularis library IL2 Bothrops insularis cDNA
            3' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401418
VERSION     BM401418.1  GI:20376046
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
```

```
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL10A (see Reference)
            Seq primer: M13R.
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL2"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      160 a    172 c    152 g    157 t     16 others
ORIGIN

  Query Match             31.8%;  Score 172.8;  DB 12;  Length 657;
  Best Local Similarity   64.8%;  Pred. No. 4.8e-36;
  Matches  351;  Conservative    0;  Mismatches  167;  Indels   24;  Gaps    6;

Qy          3 AGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTCTG 62
              | || | |  |||| ||||||  |||||   | || ||||   ||||||| ||  |||||
Db        526 AAGGGAGTTGTACAGGGTCTTNGAACAAAAGATGAACTGGNAGGATGCAGNGAAATTCTG 467

Qy         63 CACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGACTT 122
              |||  | |||  | || |||| |||  ||||||  || | |||  || |||||||| ||
Db        466 CACACAACAGCAG-ACAGGGGACATNNGGTCTCCTTCCAGAGCAGCGAAGAAGCAGATTT 408

Qy        123 TGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGACT 182
              ||||||    | | |   | |       || |       | |    ||| |||| ||||||
Db        407 TGTGGTGTCGCTCACCTCAC------CAATTTTGAGAGACAGTTTTGTNTGGACCGGACT 354

Qy        183 GAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATCAG 242
              ||| | |       ||||       ||||||  | ||||||||||||||  |  | ||||
Db        353 GAGCGATGTCTGGAAAG------GATGCAGGTTCGAGTGGAGCGATGGCAGCGACCTCAG 300

Qy        243 TTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACAGG 302
              ||  || |          |   | |    | | | |    | ||       | ||||||
Db        299 NTACAAGGACAACTACCAGTTTGTGTTTTCAGAATATGAGTGTGTCGCATCAAAGACAAA 240

Qy        303 GTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCG------A 356
              | | ||| |||  || | |  |||  | |    ||   ||| |||||||       |
Db        239 AAATAACAAATGGAGGATTATACCCTGCACAAAGTTGGAATATTTCGTCTGCGAGTTCCA 180

Qy        357 GGCATAGTCTGAAGATCCAGCTGATTGAAGTCTGGAGAAGCAAGGAAGCCCCCCCACCCCA 416
              ||| ||||||||||||| ||||  |||||||||||||| ||||||||||||||||  |
Db        179 GGCCTAGTCTGAAGATCCGGCTGTGTGAAGTCTGGAGAATCAAGGAAGCCCCCCACTC-- 122

Qy        417 TCCCCCAACCCTGCCTAGCCACAATCTCTGCTATGCACCCTTTGCTCAACGGATGCTCTC 476
              | ||| ||||      ||| |||| |||||| | | |||||||||  ||||||||||||
Db        121 -CTCCCCACCCCCGTNTGCCGCAATNTCTGCTCTTCCCCCTTTGCTGAACGGATGCTCTC 63

Qy        477 TGTAGCTGGATCTGGTGTTGCTGCTCCTGATGGGCCGGAAG--TCAATAAATTCTGCCTA 534
```

```
                ||||||||||||||| |||||||||||||||||| ||||   |||||||||||||||
Db         62 TGTAGCTGGATCTGGTTTTGCTGCTCCTGATGGGCCAGAAGGTCCAATAAATTCTGCCTA 3

Qy        535 GC 536
              |
Db          2 NC 1


RESULT 11
BM401598
LOCUS       BM401598                406 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JH4F11F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401598
VERSION     BM401598.1  GI:20376226
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL15A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT       95 a     91 c    126 g     94 t
ORIGIN

  Query Match             29.6%;  Score 161.2;  DB 12;  Length 406;
  Best Local Similarity   79.9%;  Pred. No. 5.9e-33;
  Matches  203;  Conservative    0;  Mismatches   48;  Indels    3;  Gaps    1;

Qy          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60
              ||||||| || ||||| | |||||||  | | || | ||||  |||||||||| |||
Db        156 GAAGGGCGTTCCTACAGGTTCTTCAAAACACCGAAAAACTGGGACGATGCAGAGAGGTTC 215

Qy         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120
              ||| || || ||||||| || || ||||||||||||||||||||| ||||||||||||||
Db        216 TGCTCGGAGGAGGTGAATGGCGGCCATCTGGTCTCTATCGAAAGCGCCGGAGAAGCAGAC 275

Qy        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180
```

```
Db        276 TTTGTGGCCCAGCTGGTCACTGAGAACATACAGAGCCGCGAATTGTTTGTCTGGATCGGA 335

Qy        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Db        336 ATGAAGGTTCAAGGCAAAGAAAAGCAATGCAGC---TCGTGGAGCGATGGCTCCAGCGTC 392

Qy        241 AGTTATGAGAATTG 254

Db        393 AGTTATGAGAACTG 406
```

RESULT 12
BM401605
LOCUS BM401605 410 bp mRNA linear EST 01-MAY-2002
DEFINITION JH4H07F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401605
VERSION BM401605.1 GI:20376233
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 410)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .410
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 106 a 87 c 120 g 97 t
ORIGIN

Query Match 29.3%; Score 159.6; DB 12; Length 410;
Best Local Similarity 78.0%; Pred. No. 1.6e-32;
Matches 192; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95

Db        161 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 220

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155

Db        221 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 280

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215

Db        281 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 340

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275

Db        341 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 400

Qy        276 GTGTCT 281

Db        401 GTGTTT 406
```

RESULT 13
BM401653
LOCUS BM401653 453 bp mRNA linear EST 01-MAY-2002
DEFINITION JL2D01F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401653
VERSION BM401653.1 GI:20376281
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 453)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .453
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 120 a 96 c 132 g 104 t 1 others
ORIGIN

Query Match 28.7%; Score 156.2; DB 12; Length 453;
Best Local Similarity 77.7%; Pred. No. 1.4e-31;
Matches 188; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
```

```
Db        212 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 271

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155

Db        272 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCNAACATAGGGAA 331

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215

Db        332 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 391

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAAAA 275

Db        392 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAA 451

Qy        276 GT 277

Db        452 GT 453
```

```
RESULT 14
BM401459
LOCUS       BM401459                440 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JH1D05F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401459
VERSION     BM401459.1  GI:20376087
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      113 a     96 c    130 g    101 t
ORIGIN

  Query Match             28.0%;  Score 152.2;  DB 12;  Length 440;
```

```
  Best Local Similarity   79.0%;  Pred. No. 1.7e-30;
  Matches  181;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95

Db        211 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 270

Qy         96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155

Db        271 TATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAA 330

Qy        156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215

Db        331 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 390

Qy        216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAA 264

Db        391 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGA 439
```

```
RESULT 15
BM401659
LOCUS       BM401659                456 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2F01F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401659
VERSION     BM401659.1  GI:20376287
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      117 a    100 c    131 g    104 t      4 others
ORIGIN

  Query Match             27.8%;  Score 151.2;  DB 12;  Length 456;
```

Best Local Similarity 77.9%; Pred. No. 3.2e-30;
Matches 180; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```
Qy          36 GACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTC 95
               || ||||   |||||||||| |||||| || |||||| ||| || || ||||||||||
Db         226 GAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTC 285

Qy          96 TATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTC 155
               |||| |||||   | ||||||||||||||||| |||  || | |||| ||| |||  |
Db         286 TATCNAAAGCGATGAAGAAGCAGACTTTGTGGCCCANCTGGTCGCTCCGAACATANGGAA 345

Qy         156 AGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCAT 215
               | |||||  | |||||||||||||||||||  | |||||||| || ||||||||||
Db         346 ATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTC 405

Qy         216 AGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGA 266
                |||||||||| ||||||| |||||||||||   | ||  || | |
Db         406 GAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAA 456
```

Search completed: December 8, 2003, 12:39:48
Job time : 2262 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:45:44 ; Search time 77.4 Seconds
(without alignments)
264.544 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	715	100.0	129	24	ABU08798	Deinagkistrodon ac
2	477	66.7	152	23	AAE20179	Pigmy rattlesnake
3	475	66.4	152	22	AAM51543	Snake venom blood
4	460	64.3	154	23	AAO14520	Korean adder snake
5	434	60.7	144	23	AAE20180	Pigmy rattlesnake
6	424	59.3	131	14	AAR38222	Sequence of polype
7	422.5	59.1	133	14	AAR45156	Botrocetin alpha s
8	410.5	57.4	132	23	ABB79130	Ahylysantinfarctas
9	406	56.8	132	14	AAR38226	Sequence of polype
10	404	56.5	134	14	AAR38223	Sequence of polype
11	395	55.2	130	14	AAR38228	Sequence of polype
12	387	54.1	158	23	AAE20181	Pigmy rattlesnake
13	372	52.0	132	16	AAR72235	Vipera palestinae
14	309	43.2	133	14	AAR38230	Sequence of polype
15	293	41.0	123	14	AAR38225	Sequence of polype
16	292.5	40.9	126	23	AAO20982	M9-G4 Y58A mutant
17	290	40.6	146	22	AAM51544	Snake venom blood
18	289.5	40.5	126	23	AAO20985	M12-G4 Y63A mutant
19	288.5	40.3	126	23	AAO20983	M10-G4 K61A mutant
20	288.5	40.3	126	23	AAO20986	M13-G4 R66A mutant
21	288.5	40.3	126	23	AAO20988	M15-G4 R100A mutan
22	288.5	40.3	126	23	AAO20990	M17-G4 R103A mutan
23	288.5	40.3	126	23	AAO20993	M20-G4 F108A mutan
24	287.5	40.2	126	23	AAO20981	M8-G4 D54A mutant
25	286.5	40.1	126	16	AAR71979	Snake venom derive
26	286.5	40.1	126	21	AAY85627	Snake venom derive
27	286.5	40.1	126	23	AAO20973	126-mer wild-type
28	286.5	40.1	149	16	AAR71981	Snake venom antith
29	286.5	40.1	149	21	AAY85628	Snake venom derive
30	286.5	40.1	149	23	AAO20974	149-mer wild-type
31	285.5	39.9	126	23	AAO20994	M21-G4 D54A, D101N
32	284.5	39.8	126	23	AAO20984	M11-G4 E62A mutant
33	284.5	39.8	126	23	AAO20991	M18-G4 R105A mutan
34	284.5	39.8	126	23	AAO20992	M19-G4 E106A mutan
35	284.5	39.8	126	23	AAO20996	M23-G4 D54N, D101A
36	284.5	39.8	126	23	AAO20997	M24-G4 D54A, D101A
37	283.5	39.7	126	23	AAO20989	M16-G4 D101A mutan
38	283.5	39.7	126	23	AAO20998	M25-G4 D101A, E106
39	282.5	39.5	126	23	AAO20987	M14-G4 Y67A mutant
40	282.5	39.5	126	23	AAO20995	M22-G4 D54A, D101A
41	281.5	39.4	126	23	AAO20999	M26-G4 D101A, E106
42	280.5	39.2	126	23	AAO20980	M7-G4 K20A mutant
43	274.5	38.4	127	13	AAR24426	Sequence of the pl
44	272.5	38.1	145	23	AAO14521	Korean adder snake
45	267	37.3	127	16	AAR72236	Vipera palestinae

ALIGNMENTS

```
RESULT 1
ABU08798
ID   ABU08798 standard; Protein; 129 AA.
XX
AC   ABU08798;
XX
DT   02-JUN-2003  (first entry)
XX
DE   Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
KW   Antithrombosis; alpha chain; fibrin hydrolysis; blood clot; enzyme;
KW   platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW   myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW   angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW   ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW   unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW   peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW   thrombosis.
XX
OS   Deinagkistrodon acutus.
XX
FH   Key             Location/Qualifiers
FT   Region          1..9
FT                   /note= "Antithrombosis enzyme alpha chain N-terminus not
FT                        encoded by abx93668"
FT   Region          1..29
FT                   /note= "Specifically claimed in claim 7"
FT   Misc-difference 91
FT                   /note= "Encoded by TAT"
XX
PN   US6489451-B1.
XX
```

PD 03-DEC-2002.
XX
PF 10-APR-1998; 98US-0058740.
XX
PR 10-APR-1997; 97US-043886P.
XX
PA (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI Li BX, Cheng X;
XX
DR WPI; 2003-352116/33.
DR N-PSDB; ABX93668.
XX
PT New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT preventing and treating vaso-occulusive and thromboembolic disorders,
PT including myocardial infarction, restenosis, cerebral thrombosis and
PT unstable angina -
XX
PS Claim 12; Column 11; 19pp; English.
XX
CC The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC prevents platelet aggregation. The anti-thrombosis enzyme was
CC administered to rabbits intravenously. Thrombosis was determined before
CC and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC following administration at 0.005 micro/kg and this activity was
CC increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC composition is useful for preventing and treating vaso-occulusive and
CC thromboembolic disorders, including myocardial infarction, restenosis,
CC angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC peripheral arterial occulsion, stroke. It is also useful for treating
CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC diseases. The present sequence represents the amino acid sequence of the
CC Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
SQ Sequence 129 AA;

Query Match 100.0%; Score 715; DB 24; Length 129;
Best Local Similarity 100.0%; Pred. No. 2.8e-68;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120

Qy        121 QQDPFVCEA 129
              |||||||||
Db        121 QQDPFVCEA 129
```

RESULT 2
AAE20179
ID AAE20179 standard; Protein; 152 AA.
XX
AC AAE20179;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk3.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; Zsnk3.
XX
OS Sistrurus miliarius.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /label= Signal_peptide
FT Protein 24..152
FT /label= Mature_Zsnk3_protein
FT Disulfide-bond 25..36
FT Disulfide-bond 53..150
FT Disulfide-bond 125..142
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR N-PSDB; AAD32055.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Claim 3; Page 2; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk3.
XX
SQ Sequence 152 AA;

Query Match 66.7%; Score 477; DB 23; Length 152;
Best Local Similarity 64.3%; Pred. No. 7.8e-43;
Matches 83; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||| : ||| :|||| ||| ||| ||::|  ||||||||||| || || ||: : :
Db         24 DCPSGWSSYDQHCYRVFKQLKTWDDAERFCSEQAEGGHLVSIESSEEAAFVAQLVPENRR 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               | :::||||| | |||||| :||||||:| ||||| ||| |||:   | :||| | ||
Db         84 RAILYIWIGLRVQGKEKQCSAKWSDGSSVSYENWIEAESKTCLGLQQGTNYHKWVNIYCG 143

Qy        121 QQDPFVCEA 129
               : :||||||
Db        144 EINPFVCEA 152
```

RESULT 3
AAM51543
ID AAM51543 standard; Protein; 152 AA.

```
XX
AC   AAM51543;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin A chain.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   N-PSDB; AAI71876.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 2; Page 10; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   is the A chain of halyxin.
XX
SQ   Sequence   152 AA;

  Query Match             66.4%;  Score 475;  DB 22;  Length 152;
  Best Local Similarity   63.6%;  Pred. No. 1.3e-42;
  Matches   82;  Conservative   17;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||  | |||||||||  :|   |||  :|| || ||| | |||||||| ||||| ||::: :|
Db         24 DCPSGWSSYEGHCYNIFHLFKTWAEAERFCRKQVKGAHLVSIESSEEADFVAQLVSENMK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                 |::|||||  : |:|||| :||||||:| :|||| ||| |||:  || | || | ||
Db         84 RYGIYIWIGLRVRGKKKQCSSQWSDGSSVSYQNWIEAESKTCLGLQKETEFRKWFNIYCG 143

Qy        121 QQDPFVCEA 129
                :::||||||
Db        144 ERNPFVCEA 152


RESULT 4
AAO14520
ID   AAO14520 standard; Protein; 154 AA.
XX
AC   AAO14520;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Korean adder snake venom salmorin A chain protein.
XX
KW   Korean adder; salmorin protein A chain; snake venom;
KW   fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW   thrombin binding; blood coagulation.
XX
OS   Agkistrodon halys brevicaudus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..23
FT                   /label= Signal_peptide
FT   Protein         24..154
FT                   /note= "Mature salmorin A chain protein"
XX
PN   WO200214514-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-KR01277.
XX
PR   27-JUL-2000; 2000KR-0043470.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Chung K,  Kim D,  Koh Y;
XX
DR   WPI; 2002-241907/29.
DR   N-PSDB; AAL42015.
XX
PT   New salmorin protein derived from venom of Korean adder Agkistrodon
PT   halys brevicaudus, useful for treating thrombosis by repressing
PT   fibrinogen clotting through repression of activation of prothrombin
PT   into thrombin  -
XX
PS   Claim 2; Fig 1A; 30pp; English.
XX
CC   The invention comprises the nucleotide and protein sequences of a
CC   salmorin protein derived from the venom of Korean adder. Salmorin protein
CC   is composed of an A chain and a B chain, and has inhibitory activity
CC   against fibrinogen clotting. The salmorin protein of the invention is
CC   useful for treating thrombosis, as it represses fibrinogen clotting
CC   potently by binding to prothrombin and thrombin so as to delay blood
CC   coagulation. The present amino acid sequence represents the Korean adder
CC   salmorin A chain protein.
XX
SQ   Sequence   154 AA;

  Query Match             64.3%;  Score 460;  DB 23;  Length 154;
  Best Local Similarity   60.2%;  Pred. No. 5.1e-41;
  Matches   77;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              | | | |  ||||: | |  || ||| ||: |  |||||||||:  |||||  ::|::|::
Db         27 CPSGWGSNNGHCYQAFNQRMTWEDAERFCSAQAKGGHLVSIETRAEADFVAHVVAERIET 86

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
              :  ||||||| : ||:||| |||||||:| ||||| ||| |||: ::: :||| | || |
Db         87 SFPHVWIGLRDEGKEQQCSSEWSDGSSVSYENWIEAESKTCLGLELDSNYHKWVNVYCGQ 146

Qy        122 QDPFVCEA 129
                ::||||||
Db        147 RNPFVCEA 154


RESULT 5
AAE20180
ID   AAE20180 standard; Protein; 144 AA.
XX
AC   AAE20180;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk4.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; Zsnk4.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..9
FT                   /label= Signal_peptide
```

```
FT   Protein          10..144
FT                    /label= Mature_Zsnk4_protein
FT   Disulfide-bond   13..24
FT   Disulfide-bond   41..138
FT   Disulfide-bond   113..130
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   N-PSDB; AAD32057.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 3; Page 3; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk4.
XX
SQ   Sequence   144 AA;

  Query Match             60.7%;  Score 434;  DB 23;  Length 144;
  Best Local Similarity   58.6%;  Pred. No. 2.8e-38;
  Matches   75;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||| :|: :||:| || :|| ||| ||::|  |||||||| ||| || ||:|: ||
Db         12 DCPSDWYAYDQYCYRVIKQLRTWDDAERFCSEQAKGGHLVSIESDGEAAFVAQLVAENIK 71

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                |  |||||| | :||||| :|||||:: || |:  :||| |:  ||||  | | :|
Db         72 QNKYDVWIGLRIQGEEKQCSTKWSDGSSVNYENLIKHATKKCFGLKKETGFRTWRNVHCT 131

Qy        121 QQDPFVCE 128
                ||: |:|:
Db        132 QQNLFMCK 139


RESULT 6
AAR38222
ID   AAR38222 standard; protein; 131 AA.
XX
AC   AAR38222;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin A (AL-A).
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Trimeresurus albolabris.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 4; Page 45; 74pp; English.
XX
CC   AL-A may be obtd. from snake venom. It has a mol. wt. of about 45
CC   kDa. AL-A contains two types of polypeptide chains, with mol. wt. of
CC   about 18 kDa and about 15 kDa, respectively. Each of these two types
CC   of chains is actually composed of two subtypes. Thus, AL-A comprises
CC   four non-identical polypeptide chains. A possible variation of
CC   AAR38224 has been found, characterised by Asp at posn. three in lieu
CC   of Cys.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   131 AA;

  Query Match             59.3%;  Score 424;  DB 14;  Length 131;
  Best Local Similarity   57.0%;  Pred. No. 2.8e-37;
  Matches   73;  Conservative   22;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||||||: :||:|||: :|| ||| ||::| | |||||||:|||||| ||::: |:
Db          1 DCPSDWSSYDQYCYRVFKRLQTWEDAERFCSQQANDGHLVSIESAGEADFVTQLVSENIR 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | | :|||||| | | :||| |||||||:  :|  |: ::|| |:     |  | | ||
Db         61 SEKHYVWIGLRVQGKGQQCSSEWSDGSSVHWDNLQEKNTRKCYGLEKRAEFRTWSNVYCG 120

Qy        121 QQDPFVCE 128
                : ||||:
Db        121 HEYPFVCK 128


RESULT 7
AAR45156
ID   AAR45156 standard; protein; 133 AA.
XX
AC   AAR45156;
XX
DT   25-MAR-2003  (updated)
DT   08-JUN-1994  (first entry)
XX
DE   Botrocetin alpha subunit.
XX
KW   Botrocetin; alpha subunit; snake venom; Bothrops jararaca;
KW   von Willebrand factor; von Willebrand disease.
XX
```

OS Bothrops jararaca.
XX
PN EP574621-A1.
XX
PD 22-DEC-1993.
XX
PF 02-JUN-1992; 92EP-0305046.
XX
PR 02-JUN-1992; 92EP-0305046.
XX
PA (GENZ) GENZYME CORP.
PA (UYFU-) UNIV FUJITA HEALTH.
XX
PI Fujimura Y, Titani K;
XX
DR WPI; 1993-407054/51.
XX
PT Use of purified botrocetin and opt. ristocetin - for detection of
PT von Willebrand Factor in serum or plasma, for diagnosis of von
PT Willebrand's disease
XX
PS Disclosure; Page 8; 14pp; English.
XX
CC Botrocetin, isolated from Bothrops jararaca venom, is a disulfide-
CC linked heterodimer comprising an alpha subunit (AAR45156) and a beta
CC subunit (AAR45157). The protein is used to assay von Willebrand
CC factor in serum or plasma, for diagnosis of von Willebrand disease.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 133 AA;

Query Match 59.1%; Score 422.5; DB 14; Length 133;
Best Local Similarity 56.9%; Pred. No. 4.2e-37;
Matches 74; Conservative 20; Mismatches 35; Indels 1; Gaps 1;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIE-SSGEADFVGQLIAQKI 59
              || | ||||||:||| |:|   | ||| ||::|  |||||||:  | | |||| |: : |
Db          1 DCPSGWSSYEGNCYKFFQQKMNWADAERFCSEQAKGGHLVSIKIYSKEKDFVGDLVTKNI 60

Qy         60 KSAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEEESKKCLGVHIETGFHKWENFYC 119
              :|: :: ||||| :||||||| |||||||:| || :|   |||  :  : ||  | | ||
Db         61 QSSDLYAWIGLRVENKEKQCSSEWSDGSSVSYENVVERTVKKCFALEKDLGFVLWINLYC 120

Qy        120 EQQDPFVCEA 129
                |::||||::
Db        121 AQKNPFVCKS 130
```

RESULT 8
ABB79130
ID ABB79130 standard; protein; 132 AA.
XX
AC ABB79130;
XX
DT 06-AUG-2002 (first entry)
XX
DE Ahylysantinfarctase thrombase I subunit A SEQ ID NO:1.
XX
KW Ahylysantinfarctase thrombase I; Chinese Agkistrodon acutus;
KW snake venom; enzyme; Deinagkistrodon acutus.
XX
OS Deinagkistrodon acutus.
XX
PN CN1332241-A.
XX
PD 23-JAN-2002.
XX
PF 29-APR-2001; 2001CN-0115567.
XX
PR 29-APR-2001; 2001CN-0115567.
XX
PA (KUNM-) KUNMING ZOOLOGY INST CHINESE ACAD SCI.
XX
PI Xiao C;
XX
DR WPI; 2002-281832/33.
XX
PT Ahylysantinfarctase thrombase and its production process -
XX
PS Claim 4; Page 5 (Disclosure); 8pp; Chinese.
XX
CC The present sequence represents subunit A of the ahylysantinfarctase
CC thrombase I isolated from the snake venom of Chinese Agkistrodon acutus
CC (Deinagkistrodon acutus). The present invention describes an anionic
CC exchange column chromatographic process used to separate and purify, and
CC the repurifying in a fast protein purifying work station, where
CC Ahylysantinfarctase thrombase I of over 97% purity is obtained from
CC snake venom of Chinese Agkistrodon acutus. The test shows that the
CC Ahylysantinfarctase thrombase I consists of two subunits designated A
CC and B (see ABB79130 and ABB79131 respectively).
XX
SQ Sequence 132 AA;

Query Match 57.4%; Score 410.5; DB 23; Length 132;
Best Local Similarity 58.6%; Pred. No. 7.9e-36;
Matches 75; Conservative 18; Mismatches 32; Indels 3; Gaps 2;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|   ||:|: |||   |: | | ||| |||:| :|||||||||| || ||: ||::| |:
Db          3 NCPPGWSAYD-HCY--IKEPKNWDDAEKFCTEQADGGHLVSIESKGERDFLAQLVSQSIE 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | : ||| ||| |||||||| |||||||:| || :|   :||  :  :|||||| |  |
Db         60 SIEDHVWTGLRLQNKEKQCSTEWSDGSSLSFENLLELYMRKCGALDRDTGFHKWINLGCI 119

Qy        121 QQDPFVCE 128
              | :||||:
Db        120 QLNPFVCK 127
```

RESULT 9
AAR38226
ID AAR38226 standard; protein; 132 AA.
XX
AC AAR38226;
XX
DT 25-MAR-2003 (updated)
DT 01-OCT-1993 (first entry)
XX
DE Sequence of polypeptide chain of alboaggregin B1 (AL-B1).
XX
KW Venom; snake; platelet-binding protein.
XX
OS Trimeresurus albolabris.
XX
PN WO9311151-A1.
XX
PD 10-JUN-1993.
XX
PF 01-DEC-1992; 92WO-US10344.
XX
PR 03-DEC-1991; 91US-0803630.
PR 05-JUN-1992; 92US-0893929.
XX
PA (UTEM) UNIV TEMPLE.
XX
PI Kirby EP, Peng M;
XX
DR WPI; 1993-196991/24.
XX
PT New platelet-binding proteins obtained from snake venom - inhibit
PT binding of von Willebrand factor to platelet membrane
PT glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS Claim 6; Page 47; 74pp; English.
XX

```
CC   Alboaggregin B1 and B2 (collectively "AL-B"), may be obtained in
CC   substantially pure chemical form from venom. Each of AL-B11 and -B2
CC   comprise an about 23 kDa protein formed by an about 17 kDa
CC   polypeptide chain crosslinked to an about 14 kDa polypeptide chain
CC   by one or more interchain disulfide bonds. Each of the two
CC   polypeptide chains of AL-B1 and -B2 show strong homology to two of
CC   the chains of AL-A. AL-B1 and -B2 each comprise two non-identical
CC   polypeptide chains.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   132 AA;

  Query Match             56.8%;  Score 406;  DB 14;  Length 132;
  Best Local Similarity   54.7%;  Pred. No. 2.4e-35;
  Matches   70;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||||||:: :||:: |: ||| ||| ||::| | |||||||| || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAEXFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              : | :|||||  ||||:||| |||||||:  || |:   |||  :  |: |  | | |||
Db         61 TTKYNVWIGLSVQNKEQQCSSEWSDGSSVXYENLIKPNPKKCFVLKKESEFRTWSNVYCE 120

Qy        121 QQDPFVCE 128
                 |:  |:|:
Db        121 QKHIFMCK 128


RESULT 10
AAR38223
ID   AAR38223 standard; protein; 134 AA.
XX
AC   AAR38223;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin A (AL-A).
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Trimeresurus albolabris.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 4; Page 46; 74pp; English.
XX
CC   AL-A may be obtd. from snake venom. It has a mol. wt. of about 45
CC   kDa. AL-A contains two types of polypeptide chains, with mol. wt. of
CC   about 18 kDa and about 15 kDa, respectively. Each of these two types
CC   of chains is actually composed of two subtypes. Thus, AL-A comprises
CC   four non-identical polypeptide chains. A possible variation of
CC   AAR38224 has been found, characterised by Asp at posn. three in lieu
CC   of Cys.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   134 AA;

  Query Match             56.5%;  Score 404;  DB 14;  Length 134;
  Best Local Similarity   55.9%;  Pred. No. 3.9e-35;
  Matches   71;  Conservative   22;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |   ||:|: :||:|| : | | ||| || || : |||||||: |||||| |||:: |:|
Db          4 CLPGWSAYDQYCYRVFNEPKNWEDAERFCAKQADSGHLVSIETMGEADFVAQLISENIQS 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
                : :|||||: ||||:||| |||||||:: || |:   :||  :  |:|| || |  | |
Db         64 KEHYVWIGLKVQNKEQQCSSEWSDGSSVTYENLIKLYMRKCGALEQESGFRKWINLGCIQ 123

Qy        122 QDPFVCE 128
                :||||:
Db        124 LNPFVCK 130


RESULT 11
AAR38228
ID   AAR38228 standard; protein; 130 AA.
XX
AC   AAR38228;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin B1 (AL-B2).
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Trimeresurus albolabris.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 8; Page 48; 74pp; English.
XX
CC   Alboaggregin B1 and B2 (collectively "AL-B"), may be obtained in
CC   substantially pure chemical form from venom. Each of AL-B11 and -B2
CC   comprise an about 23 kDa protein formed by an about 17 kDa
CC   polypeptide chain crosslinked to an about 14 kDa polypeptide chain
CC   by one or more interchain disulfide bonds. Each of the two
CC   polypeptide chains of AL-B1 and -B2 show strong homology to two of
CC   the chains of AL-A. AL-B1 and -B2 each comprise two non-identical
CC   polypeptide chains.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   130 AA;

  Query Match             55.2%;  Score 395;  DB 14;  Length 130;
  Best Local Similarity   54.7%;  Pred. No. 3.5e-34;
  Matches   70;  Conservative   22;  Mismatches   34;  Indels    2;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
```

```
                 || |||||:: :||:: |: ||| ||| ||::| | |||||||| || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAERFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                | |||||| ||| :||| |||||||:| || ::  |||  :  |: |  | | |||
Db         61 --KYHVWIGLSVQNKGQQCSSEWSDGSSVSYENLVKPNPKKCFVLKKESEFKTWSNVYCE 118

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        119 QKHIFMCK 126


RESULT 12
AAE20181
ID   AAE20181 standard; Protein; 158 AA.
XX
AC   AAE20181;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk5.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; Zsnk5.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..23
FT                   /label= Signal_peptide
FT   Protein         24..158
FT                   /label= Mature_Zsnk5_protein
FT   Disulfide-bond  27..38
FT   Disulfide-bond  55..152
FT   Disulfide-bond  127..144
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   N-PSDB; AAD32059.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 3; Page 3; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk5.
XX
SQ   Sequence   158 AA;

  Query Match             54.1%;  Score 387;  DB 23;  Length 158;
  Best Local Similarity   52.3%;  Pred. No. 3.1e-33;
  Matches   67;  Conservative   25;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | | :|: :||:| |: ||| ||| ||::|  |||| |:|:  || |: ||:|  ||
Db         26 NCPSGWFAYDQYCYRVIKRLKTWDDAERFCSEQAKGGHLASVENDEEAVFLAQLVAANIK 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               : :|||||| ||| :||| :||||||:| || ::  |||| |:  || | :| |  ||
Db         86 QNQYYVWIGLRIQNKGQQCSTKWSDGSSVSYENLVKSHSKKCFGLKKETEFLQWYNTDCE 145

Qy        121 QQDPFVCE 128
                ::: |||:
Db        146 EKNLFVCK 153


RESULT 13
AAR72235
ID   AAR72235 standard; peptide; 132 AA.
XX
AC   AAR72235;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide.
XX
KW   Vipera palestinae; venom; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 5; Page 19; 37pp; Japanese.
XX
CC   AAR72235 is a Vipera palestinae venom derived peptide. A peptide
CC   compsn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
CC   binding of platelets to von Willebrand factors and can therefore
CC   be used as an anti- thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   132 AA;

  Query Match             52.0%;  Score 372;  DB 16;  Length 132;
  Best Local Similarity   49.6%;  Pred. No. 9.9e-32;
```

```
  Matches   64;  Conservative   27;  Mismatches   36;  Indels    2;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:|||||||||:  ||| :|| || :|||| || |||| ||:|| :|:::  :
Db          3 DCPSDWSSHEGHCYKVFRLFKTWEEAEKFCMQQVNGWHLASIESVEEANFVAELVSKTLI 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               :| | ||||| |:: :|||  |:|||::| |      :  || |:: : |: :|    ||
Db         63 KSKYHAWIGLRDQSERQQCSSHWTDGSAVSYET--VTKYTKCFGLNKDKGYLEWVTLPCE 120

Qy        121 QQDPFVCEA 129
                : ||:|::
Db        121 DKIPFICKS 129


RESULT 14
AAR38230
ID   AAR38230 standard; protein; 133 AA.
XX
AC   AAR38230;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin-like protein called
DE   echicetin.
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Echis carinatus.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 10; Page 49; 74pp; English.
XX
CC   Echicetin, is obtained in substantially pure chemical for from
CC   venom, as an about 26 kDa protein which likely comprises a dimer
CC   composed of two types of chains. One chain type has a mol. wt. of
CC   about 15 kDa and the other has a mol. wt. of about 13 kDa. The
CC   individual polypeptide chains are held together by one or more
CC   interchain disulfide bonds.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   133 AA;

  Query Match             43.2%;  Score 309;  DB 14;  Length 133;
  Best Local Similarity   47.3%;  Pred. No. 5.1e-25;
  Matches   62;  Conservative   24;  Mismatches   39;  Indels    6;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | || |||||::|:  ||| :|| :|  | :|||||||||: :|:|| |||::|:
Db          3 DCLSGWSFYEGHCYQLFR-LKTWDEAEKYC-NQWDGGHLVSIESNAKAEFVAQLISRKLP 60

Qy         61 SAKI--HVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFY 118
               : |   |||||| ::| :||    |:| | :  |: :      ||  :  :| | ||
Db         61 KSAIEDRVWIGLRDRSKREQCGHLWTDNSFVHYEHVV--PPTKCFVLERQTEFRKWIAVN 118

Qy        119 CEQQDPFVCEA 129
                || : ||||:|
Db        119 CEFKFPFVCKA 129


RESULT 15
AAR38225
ID   AAR38225 standard; protein; 123 AA.
XX
AC   AAR38225;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin A (AL-A).
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Trimeresurus albolabris.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 4; Page 47; 74pp; English.
XX
CC   AL-A may be obtd. from snake venom. It has a mol. wt. of about 45
CC   kDa. AL-A contains two types of polypeptide chains, with mol. wt. of
CC   about 18 kDa and about 15 kDa, respectively. Each of these two types
CC   of chains is actually composed of two subtypes. Thus, AL-A comprises
CC   four non-identical polypeptide chains. A possible variation of
CC   AAR38224 has been found, characterised by Asp at posn. three in lieu
CC   of Cys.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   123 AA;

  Query Match             41.0%;  Score 293;  DB 14;  Length 123;
  Best Local Similarity   46.9%;  Pred. No. 2.4e-23;
  Matches   60;  Conservative   18;  Mismatches   42;  Indels    8;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||||||||||:|| : : | ||| |||:|  | |||| :||| |||| |:  : |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKFCTQQHKGSHLVSFQSSGGADFVVQM-TRPIL 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               :| : |||||    |     |: :||||: :  : |   |  :||      |  ::| : |
Db         60 NANL-VWIGL--SNLWNGCNSQWSDGTXLDYKXW--REQFECL--VSRTTNNEWLSMDCS 112

Qy        121 QQDPFVCE 128
                 ||||
Db        113 STHSFVCE 120


Search completed: December  8, 2003, 09:52:34
Job time : 78.4 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:50:40 ; Search time 28.0114 Seconds
(without alignments)
194.853 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	715	100.0	129	4	US-09-058-740-2	Sequence 2, Appli
2	425	59.4	131	1	US-07-893-929A-1	Sequence 1, Appli
3	425	59.4	131	5	PCT-US92-10344-1	Sequence 1, Appli
4	406	56.8	132	1	US-07-893-929A-5	Sequence 5, Appli
5	406	56.8	132	5	PCT-US92-10344-5	Sequence 5, Appli
6	404	56.5	134	1	US-07-893-929A-2	Sequence 2, Appli
7	404	56.5	134	5	PCT-US92-10344-2	Sequence 2, Appli
8	395	55.2	130	1	US-07-893-929A-7	Sequence 7, Appli
9	395	55.2	130	5	PCT-US92-10344-7	Sequence 7, Appli
10	309	43.2	133	1	US-07-893-929A-9	Sequence 9, Appli
11	309	43.2	133	5	PCT-US92-10344-9	Sequence 9, Appli
12	292	40.8	123	1	US-07-893-929A-4	Sequence 4, Appli
13	292	40.8	123	5	PCT-US92-10344-4	Sequence 4, Appli
14	286.5	40.1	126	2	US-08-612-840A-2	Sequence 2, Appli
15	286.5	40.1	149	2	US-08-612-840A-8	Sequence 8, Appli
16	274.5	38.4	127	1	US-07-614-443A-1	Sequence 1, Appli
17	274.5	38.4	127	1	US-08-294-859-1	Sequence 1, Appli
18	274.5	38.4	127	1	US-08-481-676-1	Sequence 1, Appli
19	259	36.2	117	1	US-07-614-443A-2	Sequence 2, Appli
20	259	36.2	117	1	US-08-294-859-2	Sequence 2, Appli
21	259	36.2	117	1	US-08-481-676-2	Sequence 2, Appli
22	246.5	34.5	107	1	US-07-893-929A-6	Sequence 6, Appli
23	246.5	34.5	107	5	PCT-US92-10344-6	Sequence 6, Appli
24	230.5	32.2	123	1	US-07-893-929A-10	Sequence 10, Appl
25	230.5	32.2	123	5	PCT-US92-10344-10	Sequence 10, Appl
26	222	31.0	125	1	US-07-893-929A-3	Sequence 3, Appli
27	222	31.0	125	5	PCT-US92-10344-3	Sequence 3, Appli
28	214	29.9	50	1	US-07-614-443A-5	Sequence 5, Appli
29	214	29.9	50	1	US-08-294-859-5	Sequence 5, Appli
30	214	29.9	50	1	US-08-481-676-5	Sequence 5, Appli
31	214	29.9	128	1	US-07-893-929A-8	Sequence 8, Appli
32	214	29.9	128	5	PCT-US92-10344-8	Sequence 8, Appli
33	201.5	28.2	1479	3	US-08-840-062-4	Sequence 4, Appli
34	197.5	27.6	1479	3	US-08-840-062-2	Sequence 2, Appli
35	192.5	26.9	166	2	US-08-729-103-4	Sequence 4, Appli
36	186.5	26.1	174	2	US-08-401-530A-5	Sequence 5, Appli
37	186.5	26.1	174	2	US-08-709-662-5	Sequence 5, Appli
38	184.5	25.8	1257	1	US-08-340-428B-49	Sequence 49, Appl
39	183	25.6	174	1	US-07-778-156-7	Sequence 7, Appli
40	183	25.6	174	2	US-08-822-261-4	Sequence 4, Appli
41	183	25.6	174	2	US-08-422-166-7	Sequence 7, Appli
42	183	25.6	174	4	US-09-226-852-4	Sequence 4, Appli
43	183	25.6	175	2	US-08-464-637-2	Sequence 2, Appli
44	183	25.6	175	2	US-08-401-530A-4	Sequence 4, Appli
45	183	25.6	175	2	US-08-709-662-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-058-740-2
; Sequence 2, Application US/09058740
; Patent No. 6489451
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,740
; FILING DATE: 10-Apr-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 129 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-058-740-2

Query Match 100.0%; Score 715; DB 4; Length 129;
Best Local Similarity 100.0%; Pred. No. 7.1e-71;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120

Qy        121 QQDPFVCEA 129
              |||||||||
Db        121 QQDPFVCEA 129
```

RESULT 2
US-07-893-929A-1
; Sequence 1, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-893-929A-1

Query Match 59.4%; Score 425; DB 1; Length 131;
Best Local Similarity 57.0%; Pred. No. 4e-39;
Matches 73; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||: :||:|||: :|| ||| ||::| | ||||||||:|||||| ||::: |:
Db          1 DCPSDWSSYDQYCYRVFKRLQTWEDAERFCSEQANDGHLVSIESAGEADFVTQLVSENIR 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | | :|||||| | | :||| |||||||: :|  |: ::|| |:     |  | | ||
Db         61 SEKHYVWIGLRVQGKGQQCSSEWSDGSSVHYDNLQEKNTRKCYGLEKRAEFRTWSNVYCG 120

Qy        121 QQDPFVCE 128
              : ||||:
Db        121 HEYPFVCK 128
```

RESULT 3
PCT-US92-10344-1
; Sequence 1, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-10344-1

Query Match 59.4%; Score 425; DB 5; Length 131;
Best Local Similarity 57.0%; Pred. No. 4e-39;
Matches 73; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||: :||:|||: :|| ||| ||::| | ||||||||:|||||| ||::: |:
Db          1 DCPSDWSSYDQYCYRVFKRLQTWEDAERFCSEQANDGHLVSIESAGEADFVTQLVSENIR 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | | :|||||| | | :||| |||||||: :|  |: ::|| |:     |  | | ||
Db         61 SEKHYVWIGLRVQGKGQQCSSEWSDGSSVHYDNLQEKNTRKCYGLEKRAEFRTWSNVYCG 120

Qy        121 QQDPFVCE 128
              : ||||:
Db        121 HEYPFVCK 128
```

RESULT 4
US-07-893-929A-5

```
; Sequence 5, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  132 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-893-929A-5

  Query Match             56.8%;  Score 406;  DB 1;  Length 132;
  Best Local Similarity   54.7%;  Pred. No. 4.8e-37;
  Matches   70;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:: :||:: |: ||| ||| ||::| | |||||||  || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAEKFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              : | :|||||  ||||:||| |||||||:  || |:   |||  :  |: |  | | |||
Db         61 TTKYNVWIGLSVQNKEQQCSSEWSDGSSVXYENLIKPNPKKCFVLKKESEFRTWSNVYCE 120

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        121 QKHIFMCK 128


RESULT 5
PCT-US92-10344-5
; Sequence 5, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10344
;      FILING DATE:  19921201
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  None
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  132 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-10344-5

  Query Match             56.8%;  Score 406;  DB 5;  Length 132;
  Best Local Similarity   54.7%;  Pred. No. 4.8e-37;
  Matches   70;  Conservative   23;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:: :||:: |: ||| ||| ||::| | |||||||  || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAEKFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              : | :|||||  ||||:||| |||||||:  || |:   |||  :  |: |  | | |||
Db         61 TTKYNVWIGLSVQNKEQQCSSEWSDGSSVXYENLIKPNPKKCFVLKKESEFRTWSNVYCE 120

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        121 QKHIFMCK 128


RESULT 6
US-07-893-929A-2
; Sequence 2, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
```

```
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  134 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-893-929A-2

  Query Match             56.5%;  Score 404;  DB 1;  Length 134;
  Best Local Similarity   55.9%;  Pred. No. 8.1e-37;
  Matches   71;  Conservative   22;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |   ||:|: :||:|| : | | ||| || || : |||||||: |||||| |||:: |:|
Db          4 CLPGWSAYDQYCYRVFNEPKNWEDAERFCAKQADSGHLVSIETMGEADFVAQLISENIQS 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
               : :|||||: ||||:||| |||||||:: || |:   :|| :  |:|| || |  | |
Db         64 KEHYVWIGLKVQNKEQQCSSEWSDGSSVTYENLIKLYMRKCGALEQESGFRKWINLGCIQ 123

Qy        122 QDPFVCE 128
               :||||:
Db        124 LNPFVCK 130


RESULT 7
PCT-US92-10344-2
; Sequence 2, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10344
;      FILING DATE:  19921201
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  None
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  134 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-10344-2

  Query Match             56.5%;  Score 404;  DB 5;  Length 134;
  Best Local Similarity   55.9%;  Pred. No. 8.1e-37;
  Matches   71;  Conservative   22;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |   ||:|: :||:|| : | | ||| || || : |||||||: |||||| |||:: |:|
Db          4 CLPGWSAYDQYCYRVFNEPKNWEDAERFCAKQADSGHLVSIETMGEADFVAQLISENIQS 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
               : :|||||: ||||:||| |||||||:: || |:   :|| :  |:|| || |  | |
Db         64 KEHYVWIGLKVQNKEQQCSSEWSDGSSVTYENLIKLYMRKCGALEQESGFRKWINLGCIQ 123

Qy        122 QDPFVCE 128
               :||||:
Db        124 LNPFVCK 130


RESULT 8
US-07-893-929A-7
; Sequence 7, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  130 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
```

US-07-893-929A-7

Query Match 55.2%; Score 395; DB 1; Length 130;
Best Local Similarity 54.7%; Pred. No. 7.6e-36;
Matches 70; Conservative 22; Mismatches 34; Indels 2; Gaps 1;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:: :||:: |: ||| ||| ||::| | |||||||  || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAERFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                | ||||||  ||| :||| |||||||:| || ::   |||  :  |: |  | | |||
Db         61 --KYHVWIGLSVQNKGQQCSSEWSDGSSVSYENLVKPNPKKCFVLKKESEFKTWSNVYCE 118

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        119 QKHIFMCK 126
```

RESULT 9
PCT-US92-10344-7
; Sequence 7, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-10344-7

Query Match 55.2%; Score 395; DB 5; Length 130;
Best Local Similarity 54.7%; Pred. No. 7.6e-36;
Matches 70; Conservative 22; Mismatches 34; Indels 2; Gaps 1;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:: :||:: |: ||| ||| ||::| | |||||||  || || :|::: :|
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAERFCSEQANDGHLVSIESYREAVFVAELLSENVK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                | ||||||  ||| :||| |||||||:| || ::   |||  :  |: |  | | |||
Db         61 --KYHVWIGLSVQNKGQQCSSEWSDGSSVSYENLVKPNPKKCFVLKKESEFKTWSNVYCE 118

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        119 QKHIFMCK 126
```

RESULT 10
US-07-893-929A-9
; Sequence 9, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-893-929A-9

Query Match 43.2%; Score 309; DB 1; Length 133;
Best Local Similarity 47.3%; Pred. No. 2e-26;
Matches 62; Conservative 24; Mismatches 39; Indels 6; Gaps 4;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | || |||||::|:  ||| :|| :|  | :|||||||||: :|:|| |||::|:
Db          3 DCLSGWSFYEGHCYQLFR-LKTWDEAEKYC-NQWDGGHLVSIESNAKAEFVAQLISRKLP 60

Qy         61 SAKI--HVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFY 118
                : |   |||||| ::| :||   |:| | :  |: :    ||  :  :| | ||
Db         61 KSAIEDRVWIGLRDRSKREQCGHLWTDNSFVHYEHVV--PPTKCFVLERQTEFRKWIAVN 118

Qy        119 CEQQDPFVCEA 129
                || : ||||:|
Db        119 CEFKFPFVCKA 129
```

RESULT 11
PCT-US92-10344-9

```
; Sequence 9, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10344
;      FILING DATE:  19921201
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  None
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  133 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-10344-9

  Query Match             43.2%;  Score 309;  DB 5;  Length 133;
  Best Local Similarity   47.3%;  Pred. No. 2e-26;
  Matches   62;  Conservative   24;  Mismatches   39;  Indels    6;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | || ||||||::|:  ||| :|| :|  | :|||||||||: :|:|| |||::|:
Db          3 DCLSGWSFYEGHCYQLFR-LKTWDEAEKYC-NQWDGGHLVSIESNAKAEFVAQLISRKLP 60

Qy         61 SAKI--HVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFY 118
                :|   |||||| ::| :||    |:| | :  |: :     ||  :  :| | ||
Db         61 KSAIEDRVWIGLRDRSKREQCGHLWTDNSFVHYEHVV--PPTKCFVLERQTEFRKWIAVN 118

Qy        119 CEQQDPFVCEA 129
              || : ||||:|
Db        119 CEFKFPFVCKA 129


RESULT 12
US-07-893-929A-4
; Sequence 4, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  123 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-893-929A-4

  Query Match             40.8%;  Score 292;  DB 1;  Length 123;
  Best Local Similarity   46.9%;  Pred. No. 1.3e-24;
  Matches   60;  Conservative   18;  Mismatches   42;  Indels    8;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||:|| : : | ||| |||:|  | |||| :|| ||||| |:  : |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKFCTQQHKGSHLVSFQSSEEADFVVQM-TRPIL 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              :| : |||||    |     |: :||||: :  : |   |  :||      |  ::| :  |
Db         60 NANL-VWIGL--SNLWNGCNSQWSDGTXLDYKXW--REQFECL--VSRTTNNEWLSMDCS 112

Qy        121 QQDPFVCE 128
                 ||||
Db        113 STHSFVCE 120


RESULT 13
PCT-US92-10344-4
; Sequence 4, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10344
;      FILING DATE:  19921201
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  None
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  123 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
PCT-US92-10344-4

  Query Match             40.8%;  Score 292;  DB 5;  Length 123;
  Best Local Similarity   46.9%;  Pred. No. 1.3e-24;
  Matches   60;  Conservative   18;  Mismatches   42;  Indels    8;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||:|| : : | ||| |||:|  | |||| :|| ||||| |:  : |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKFCTQQHKGSHLVSFQSSEEADFVVQM-TRPIL 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              :| : |||||    |    |: :||||: :  : |   | :||      |  ::| :  |
Db         60 NANL-VWIGL--SNLWNGCNSQWSDGTXLDYKXW--REQFECL--VSRTTNNEWLSMDCS 112

Qy        121 QQDPFVCE 128
                 ||||
Db        113 STHSFVCE 120


RESULT 14
US-08-612-840A-2
; Sequence 2, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  126 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-612-840A-2

  Query Match             40.1%;  Score 286.5;  DB 2;  Length 126;
  Best Local Similarity   43.0%;  Pred. No. 5.5e-24;
  Matches   55;  Conservative   22;  Mismatches   42;  Indels    9;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | ||||: :||| ||| || ||| ||::|  ||||:|:|:: || ||  ::    :
Db          3 ECPSGWSSYDRYCYKPFKQEMTWADAERFCSEQAKGGHLLSVETALEASFVDNVLYANKE 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| ||| : |       |||| || :  :  :|  |  :|   :|    ||
Db         63 YLTRYIWIGLRVQNKGQPC-------SSISYENLV--DPFECFMVSRDTRLREWFKVDCE 113

Qy        121 QQDPFVCE 128
                 ||  |:|:
Db        114 QQHSFICK 121


RESULT 15
US-08-612-840A-8
; Sequence 8, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
```

```
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-612-840A-8

  Query Match             40.1%;  Score 286.5;  DB 2;  Length 149;
  Best Local Similarity   43.0%;  Pred. No. 6.8e-24;
  Matches   55;  Conservative   22;  Mismatches   42;  Indels    9;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | ||||: :||| |||  || ||| ||::|  ||||:|:|:: || ||  ::    :
Db         26 ECPSGWSSYDRYCYKPFKQEMTWADAERFCSEQAKGGHLLSVETALEASFVDNVLYANKE 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| ||| : |       |||| || :  :  :|  |  :|   :|    ||
Db         86 YLTRYIWIGLRVQNKGQPC-------SSISYENLV--DPFECFMVSRDTRLREWFKVDCE 136

Qy        121 QQDPFVCE 128
               ||  |:|:
Db        137 QQHSFICK 144


Search completed: December  8, 2003, 09:56:25
Job time : 29.0114 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:54:55 ; Search time 53.0743 Seconds
(without alignments)
452.044 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	715	100.0	129	11	US-09-938-114-2	Sequence 2, Appli
2	477	66.7	152	10	US-09-929-230-5	Sequence 5, Appli
3	477	66.7	152	12	US-10-226-420-5	Sequence 5, Appli
4	434	60.7	144	10	US-09-929-230-8	Sequence 8, Appli
5	434	60.7	144	12	US-10-226-420-8	Sequence 8, Appli
6	387	54.1	158	10	US-09-929-230-11	Sequence 11, Appl
7	387	54.1	158	12	US-10-226-420-11	Sequence 11, Appl
8	286.5	40.1	149	10	US-09-969-763-3	Sequence 3, Appli
9	271.5	38.0	126	10	US-09-969-763-1	Sequence 1, Appli
10	264	36.9	151	10	US-09-929-230-2	Sequence 2, Appli
11	264	36.9	151	12	US-10-226-420-2	Sequence 2, Appli
12	223.5	31.3	110	10	US-09-969-763-10	Sequence 10, Appl
13	201.5	28.2	154	9	US-09-764-870-439	Sequence 439, App
14	201.5	28.2	154	15	US-10-125-540-439	Sequence 439, App
15	201.5	28.2	166	12	US-09-997-003-30	Sequence 30, Appl
16	201.5	28.2	166	12	US-09-997-003-43	Sequence 43, Appl
17	201.5	28.2	174	9	US-09-925-301-1182	Sequence 1182, Ap
18	201.5	28.2	250	9	US-09-764-870-280	Sequence 280, App
19	201.5	28.2	250	15	US-10-125-540-280	Sequence 280, App
20	201.5	28.2	1479	15	US-10-015-219-1738	Sequence 1738, Ap
21	201.5	28.2	1479	15	US-10-015-219-1739	Sequence 1739, Ap
22	192.5	26.9	174	9	US-09-925-297-753	Sequence 753, App
23	184.5	25.8	1321	12	US-10-241-220-82	Sequence 82, Appl
24	183	25.6	174	15	US-10-316-761-4	Sequence 4, Appli
25	183	25.6	175	15	US-10-316-761-3	Sequence 3, Appli
26	183	25.6	183	9	US-09-925-301-1013	Sequence 1013, Ap
27	181.5	25.4	158	9	US-09-922-217-1070	Sequence 1070, Ap
28	181.5	25.4	158	9	US-09-922-217-1077	Sequence 1077, Ap
29	181.5	25.4	158	9	US-09-922-217-1078	Sequence 1078, Ap
30	181.5	25.4	158	9	US-09-922-217-1079	Sequence 1079, Ap
31	181.5	25.4	158	9	US-09-922-217-1080	Sequence 1080, Ap
32	181.5	25.4	158	10	US-09-833-263-1070	Sequence 1070, Ap
33	181.5	25.4	158	10	US-09-833-263-1077	Sequence 1077, Ap
34	181.5	25.4	158	10	US-09-833-263-1078	Sequence 1078, Ap
35	181.5	25.4	158	10	US-09-833-263-1079	Sequence 1079, Ap
36	181.5	25.4	158	10	US-09-833-263-1080	Sequence 1080, Ap
37	181.5	25.4	158	12	US-09-525-041-2	Sequence 2, Appli
38	181.5	25.4	158	14	US-10-025-380-1070	Sequence 1070, Ap
39	181.5	25.4	158	14	US-10-025-380-1077	Sequence 1077, Ap
40	181.5	25.4	158	14	US-10-025-380-1078	Sequence 1078, Ap
41	181.5	25.4	158	14	US-10-025-380-1079	Sequence 1079, Ap
42	181.5	25.4	158	14	US-10-025-380-1080	Sequence 1080, Ap
43	181.5	25.4	158	15	US-10-100-608B-2	Sequence 2, Appli
44	181.5	25.4	158	15	US-10-157-031-114	Sequence 114, App
45	181.5	25.4	158	15	US-10-205-823-347	Sequence 347, App

ALIGNMENTS

RESULT 1
US-09-938-114-2
; Sequence 2, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,114
; FILING DATE: 23-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,740
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298

```
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 129 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: Peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-938-114-2

  Query Match             100.0%;  Score 715;  DB 11;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-70;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120

Qy        121 QQDPFVCEA 129
              |||||||||
Db        121 QQDPFVCEA 129


RESULT 2
US-09-929-230-5
; Sequence 5, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-5

  Query Match             66.7%;  Score 477;  DB 10;  Length 152;
  Best Local Similarity   64.3%;  Pred. No. 7.1e-44;
  Matches   83;  Conservative   16;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | ||||: |||:|||| ||| ||| ||::|  |||||||||| || || ||: :  :
Db         24 DCPSGWSSYDQHCYRVFKQLKTWDDAERFCSEQAEGGHLVSIESSEEAAFVAQLVPENRR 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | :::||||| | |||||| :||||||:| ||||| ||| |||:    | :||| | ||
Db         84 RAILYIWIGLRVQGKEKQCSAKWSDGSSVSYENWIEAESKTCLGLQQGTNYHKWVNIYCG 143

Qy        121 QQDPFVCEA 129
              : :||||||
Db        144 EINPFVCEA 152


RESULT 3
US-10-226-420-5
; Sequence 5, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-5

  Query Match             66.7%;  Score 477;  DB 12;  Length 152;
  Best Local Similarity   64.3%;  Pred. No. 7.1e-44;
  Matches   83;  Conservative   16;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | ||||: |||:|||| ||| ||| ||::|  |||||||||| || || ||: :  :
Db         24 DCPSGWSSYDQHCYRVFKQLKTWDDAERFCSEQAEGGHLVSIESSEEAAFVAQLVPENRR 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | :::||||| | |||||| :||||||:| ||||| ||| |||:    | :||| | ||
Db         84 RAILYIWIGLRVQGKEKQCSAKWSDGSSVSYENWIEAESKTCLGLQQGTNYHKWVNIYCG 143

Qy        121 QQDPFVCEA 129
              : :||||||
Db        144 EINPFVCEA 152


RESULT 4
US-09-929-230-8
; Sequence 8, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 144
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-8

  Query Match             60.7%;  Score 434;  DB 10;  Length 144;
  Best Local Similarity   58.6%;  Pred. No. 3e-39;
  Matches   75;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||| :|: :||:| || :|| ||| ||::|  ||||||||| ||| || ||:|: ||
Db         12 DCPSDWYAYDQYCYRVIKQLRTWDDAERFCSEQAKGGHLVSIESDGEAAFVAQLVAENIK 71

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              |  |||||| | :||||| :||||||:: || |:  :||| |:  ||||  | | :|
Db         72 QNKYDVWIGLRIQGEEKQCSTKWSDGSSVNYENLIKHATKKCFGLKKETGFRTWRNVHCT 131

Qy        121 QQDPFVCE 128
              ||: |:|:
Db        132 QQNLFMCK 139


RESULT 5
US-10-226-420-8
; Sequence 8, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
```

```
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 144
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-8

  Query Match             60.7%;  Score 434;  DB 12;  Length 144;
  Best Local Similarity   58.6%;  Pred. No. 3e-39;
  Matches   75;  Conservative   20;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || ||| :|: :||:| || :|| ||| ||::|  ||||||||| ||| || ||:|: ||
Db         12 DCPSDWYAYDQYCYRVIKQLRTWDDAERFCSEQAKGGHLVSIESDGEAAFVAQLVAENIK 71

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               |  |||||| | :||||| :|||||:: || |:  :||| |:  ||||  | | :|
Db         72 QNKYDVWIGLRIQGEEKQCSTKWSDGSSVNYENLIKHATKKCFGLKKETGFRTWRNVHCT 131

Qy        121 QQDPFVCE 128
                ||: |:|:
Db        132 QQNLFMCK 139


RESULT 6
US-09-929-230-11
; Sequence 11, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-11

  Query Match             54.1%;  Score 387;  DB 10;  Length 158;
  Best Local Similarity   52.3%;  Pred. No. 4.2e-34;
  Matches   67;  Conservative   25;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | | :|: :||:| |: ||| ||| ||::|  |||| |:|:  || |: ||:|  ||
Db         26 NCPSGWFAYDQYCYRVIKRLKTWDDAERFCSEQAKGGHLASVENDEEAVFLAQLVAANIK 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                : :|||||| ||| :||| :|||||:| || ::  |||| |:  || | :| |  ||
Db         86 QNQYYVWIGLRIQNKGQQCSTKWSDGSSVSYENLVKSHSKKCFGLKKETEFLQWYNTDCE 145

Qy        121 QQDPFVCE 128
                ::: |||:
Db        146 EKNLFVCK 153


RESULT 7
US-10-226-420-11
; Sequence 11, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-11

  Query Match             54.1%;  Score 387;  DB 12;  Length 158;
  Best Local Similarity   52.3%;  Pred. No. 4.2e-34;
  Matches   67;  Conservative   25;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | | :|: :||:| |: ||| ||| ||::|  |||| |:|:  || |: ||:|  ||
Db         26 NCPSGWFAYDQYCYRVIKRLKTWDDAERFCSEQAKGGHLASVENDEEAVFLAQLVAANIK 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                : :|||||| ||| :||| :|||||:| || ::  |||| |:  || | :| |  ||
Db         86 QNQYYVWIGLRIQNKGQQCSTKWSDGSSVSYENLVKSHSKKCFGLKKETEFLQWYNTDCE 145

Qy        121 QQDPFVCE 128
                ::: |||:
Db        146 EKNLFVCK 153


RESULT 8
US-09-969-763-3
; Sequence 3, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING TH
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Crotalus harridus
US-09-969-763-3

  Query Match             40.1%;  Score 286.5;  DB 10;  Length 149;
  Best Local Similarity   43.0%;  Pred. No. 3.1e-23;
  Matches   55;  Conservative   22;  Mismatches   42;  Indels    9;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | |||| : :||| |||  || ||| ||::|  ||||:|:|:: || ||  ::    :
Db         26 ECPSGWSSYDRYCYKPFKQEMTWADAERFCSEQAKGGHLLSVETALEASFVDNVLYANKE 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| ||| : |       |||| || :  :  :|  |  :|   :|    ||
Db         86 YLTRYIWIGLRVQNKGQPC-------SSISYENLV--DPFECFMVSRDTRLREWFKVDCE 136

Qy        121 QQDPFVCE 128
                ||  |:|:
Db        137 QQHSFICK 144
```

```
RESULT 9
US-09-969-763-1
; Sequence 1, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 126
;   TYPE: PRT
;   ORGANISM: Crotallus horridus
US-09-969-763-1

  Query Match             38.0%;  Score 271.5;  DB 10;  Length 126;
  Best Local Similarity   41.4%;  Pred. No. 1e-21;
  Matches   53;  Conservative   22;  Mismatches   44;  Indels    9;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | ||| : :||| |||  ||  || ||::|  ||||:|:|:: || ||  ::    :
Db          3 ECPSGWSSTDRYCYKPFKQEMTWASAERFCSEQAKGGHLLSVETALEASFVDNVLYANKE 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::|||| ||| : |       |||| || :  :  :|  |  :|   :|    ||
Db         63 YLTRYIWIGLRVQNKGQPC-------SSISYENLV--DPFECFMVSRDTRLREWFKVDCE 113

Qy        121 QQDPFVCE 128
              ||  |:|:
Db        114 QQHSFICK 121


RESULT 10
US-09-929-230-2
; Sequence 2, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-2

  Query Match             36.9%;  Score 264;  DB 10;  Length 151;
  Best Local Similarity   43.0%;  Pred. No. 8.5e-21;
  Matches   55;  Conservative   18;  Mismatches   49;  Indels    6;  Gaps    3;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||| |||||| : ||| ||||||  |     | || || |  |||:| :| :|
Db         20 DCPSDWSSYDQHCYKVFSELKTWDDAESFCYTQHRDSRLASIHSSEEEAFVGKLASQTLK 79

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                :||||:  :  |:|  :||| : :  : |       |  : ::|    | |  ||
Db         80 FTS--MWIGLK--DLWKECKWQWSDDTKLDYKAWTRR--PYCTVMVVKTDRIFWFNRGCE 133

Qy        121 QQDPFVCE 128
              :  |||:
Db        134 KTVSFVCK 141


RESULT 11
US-10-226-420-2
; Sequence 2, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-2

  Query Match             36.9%;  Score 264;  DB 12;  Length 151;
  Best Local Similarity   43.0%;  Pred. No. 8.5e-21;
  Matches   55;  Conservative   18;  Mismatches   49;  Indels    6;  Gaps    3;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||| |||||| : ||| ||||||  |     | || || |  |||:| :| :|
Db         20 DCPSDWSSYDQHCYKVFSELKTWDDAESFCYTQHRDSRLASIHSSEEEAFVGKLASQTLK 79

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                :||||:  :  |:|  :||| : :  : |       |  : ::|    | |  ||
Db         80 FTS--MWIGLK--DLWKECKWQWSDDTKLDYKAWTRR--PYCTVMVVKTDRIFWFNRGCE 133

Qy        121 QQDPFVCE 128
              :  |||:
Db        134 KTVSFVCK 141


RESULT 12
US-09-969-763-10
; Sequence 10, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING TH
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 110
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-969-763-10
```

```
  Query Match             31.3%;  Score 223.5;  DB 10;  Length 110;
  Best Local Similarity   35.2%;  Pred. No. 1.4e-16;
  Matches   45;  Conservative   18;  Mismatches   40;  Indels   25;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | ||||  :||| |||  |: ||| ||::|  ||||:|:|:: || ||  ::    :
Db          3 ECPSGWSSYSRYCYKPFKQEMTYADAERFCSEQAKGGHLLSVETALEASFVDNVLYANKE 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::|||||    |                    |  | :|  :|   ||
Db         63 YLTRYIWIGLRFFFFE----------------------CFMVSRDTRLREWFKVDCE 97

Qy        121 QQDPFVCE 128
              ||  |:|:
Db         98 QQHSFICK 105


RESULT 13
US-09-764-870-439
; Sequence 439, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14
;  CURRENT APPLICATION NUMBER: US/09/764,870
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (141)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-439

  Query Match             28.2%;  Score 201.5;  DB 9;  Length 154;
  Best Local Similarity   28.9%;  Pred. No. 5.2e-14;
  Matches   39;  Conservative   32;  Mismatches   47;  Indels   17;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|   |  ::||||::  : ::| :::  | :   || |||| |  | :|: : | |:::
Db         11 ECEPSWQPFQGHCYRLQAEKRSWQESKKACLR--GGGDLVSIHSMAELEFITKQIKQEVE 68

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENW-------IEEESKKCLGVHIETGFHK 113
                :||||   : : | : ||||||  :|  :|         :  : |: :    |  :
Db         69 ----ELWIGL--NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVTIWGPEG--R 120

Qy        114 WENFYCEQQDPFVCE 128
              | :  | |  | :|:
Db        121 WNDSPCNQSLPSICK 135


RESULT 14
US-10-125-540-439
; Sequence 439, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ14C1
;  CURRENT APPLICATION NUMBER: US/10/125,540
;  CURRENT FILING DATE:  2002-04-19
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 646
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 439
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (141)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-439

  Query Match             28.2%;  Score 201.5;  DB 15;  Length 154;
  Best Local Similarity   28.9%;  Pred. No. 5.2e-14;
  Matches   39;  Conservative   32;  Mismatches   47;  Indels   17;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|   |  ::||||::  : ::| :::  | :   || |||| |  | :|: : | |:::
Db         11 ECEPSWQPFQGHCYRLQAEKRSWQESKKACLR--GGGDLVSIHSMAELEFITKQIKQEVE 68

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENW-------IEEESKKCLGVHIETGFHK 113
                :||||   : : | : ||||||  :|  :|         :  : |: :    |  :
Db         69 ----ELWIGL--NDLKLQMNFEWSDGSLVSFTHWHPFEPNNFRDSLEDCVTIWGPEG--R 120

Qy        114 WENFYCEQQDPFVCE 128
              | :  | |  | :|:
Db        121 WNDSPCNQSLPSICK 135


RESULT 15
US-09-997-003-30
; Sequence 30, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA003P1
;  CURRENT APPLICATION NUMBER: US/09/997,003
;  CURRENT FILING DATE:  2001-11-30
;  PRIOR APPLICATION NUMBER: unassigned
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: PCT/US00/22157
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/148,680
;  PRIOR FILING DATE: 1999-08-13
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-997-003-30

  Query Match             28.2%;  Score 201.5;  DB 12;  Length 166;
  Best Local Similarity   33.3%;  Pred. No. 5.7e-14;
  Matches   44;  Conservative   26;  Mismatches   53;  Indels    9;  Gaps    5;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |    ::|  :||   :  :|| ||: :| : :| |:|||: :  |  ||  ||  :  :
Db         36 CPEGTNAYRSYCYYFNEDRETWVDADLYC-QNMNSGNLVSVLTQAEGAFVASLIKES-GT 93

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENW-IEEESK----KCLGVHIETGFHKWEN 116
                :|||||    | ::    || || :| ::| |   |       |: :   ||| ||::
Db         94 DDFNVWIGLHDPKKNRRW--HWSSGSLVSYKSWGIGAPSSVNPGYCVSLTSSTGFQKWKD 151

Qy        117 FYCEQQDPFVCE 128
                || :  |||:
Db        152 VPCEDKFSFVCK 163


Search completed: December  8, 2003, 10:07:06
Job time : 54.0743 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:41:24 ; Search time 5475.5 Seconds
(without alignments)
963.811 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22849/app_query.fasta_1.725
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_3707_@runat_08122003_091000_22849 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	701	98.0	733	5	AF176420	AF176420 Deinagkis
2	657	91.9	497	5	AY091758	AY091758 Deinagkis
3	657	91.9	656	5	AB036880	AB036880 Deinagkis
4	653	91.3	544	6	AR259035	AR259035 Sequence
5	496	69.4	678	5	TFLFIXA	D83331 Trimeresuru
6	496	69.4	683	5	AB046491	AB046491 Trimeresu
7	477	66.7	721	6	AX427204	AX427204 Sequence
8	475	66.4	630	5	AF190827	AF190827 Gloydius
9	460	64.3	632	5	AF125309	AF125309 Gloydius
10	439	61.4	508	5	AF463522	AF463522 Deinagkis
11	439	61.4	661	5	AF540646	AF540646 Deinagkis
12	434	60.7	580	6	AX427207	AX427207 Sequence
13	420.5	58.8	408	5	AF244900	AF244900 Callosela
14	397	55.5	456	6	AX427206	AX427206 Sequence
15	394	55.1	477	5	AY099321	AY099321 Protoboth
16	391	54.7	681	5	AY149341	AY149341 Trimeresu
17	387	54.1	725	6	AX427210	AX427210 Sequence
18	380.5	53.2	495	5	AY091762	AY091762 Deinagkis
19	378	52.9	432	6	AX427209	AX427209 Sequence
20	377.5	52.8	704	5	AF102901	AF102901 Deinagkis
21	375	52.4	508	5	AY091759	AY091759 Deinagkis
22	375	52.4	666	5	AF540645	AF540645 Deinagkis
23	364	50.9	800	5	CDRNACVXA	Y16348 Crotalus du
24	355.5	49.7	679	5	AB019615	AB019615 Agkistrod
25	329	46.0	474	6	AX427212	AX427212 Sequence
26	291	40.7	698	5	TFLFIXB	D83332 Trimeresuru
27	290	40.6	634	5	AF197915	AF197915 Gloydius
28	290	40.6	688	5	AF540647	AF540647 Deinagkis
29	289	40.4	478	5	AY091761	AY091761 Deinagkis
30	286.5	40.1	690	6	AR026653	AR026653 Sequence
31	286.5	40.1	690	6	AX404807	AX404807 Sequence
32	286.5	40.1	690	6	BD143687	BD143687 Protein h
33	282	39.4	369	5	AY293866	AY293866 Deinagkis
34	282	39.4	574	5	AF350324	AF350324 Deinagkis
35	280	39.2	438	5	AF244901	AF244901 Callosela
36	278	38.9	458	5	AF176421	AF176421 Deinagkis
37	277	38.7	592	5	AB036881	AB036881 Deinagkis
38	274	38.3	478	5	AY091756	AY091756 Deinagkis
39	273	38.2	664	5	AB019616	AB019616 Agkistrod
40	272.5	38.1	583	5	AF125310	AF125310 Gloydius
41	271.5	38.0	487	5	AF463521	AF463521 Deinagkis
42	268.5	37.6	619	5	AF540648	AF540648 Deinagkis
43	264	36.9	483	6	AX427201	AX427201 Sequence
44	264	36.9	738	5	CDRNACVXB	Y16349 Crotalus du
45	256	35.8	484	5	AY091760	AY091760 Deinagkis

ALIGNMENTS

RESULT 1

```
AF176420
LOCUS       AF176420                733 bp    mRNA    linear   VRT 30-APR-2002
DEFINITION  Deinagkistrodon acutus clone 2100490 agkisacutacin A chain mRNA,
            complete cds.
ACCESSION   AF176420 AY091757
VERSION     AF176420.2  GI:20273043
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J. and Liu,J.
  TITLE     cDNA sequencing and analysis of eleven C-type lectin-like protein
            subunits from Agkistrodon acutus
  JOURNAL   Sheng wu hua hsueh yu sheng wu wu li hsueh pao (2002) In press
REFERENCE   2  (bases 1 to 733)
  AUTHORS   Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and
            Liu,J.
  TITLE     Agkisacutacin, a new fibrinolytic & anti-platelet protein from
            Agkistrodon acutus venom
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     A chain of agkisacutacin from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 733)
  AUTHORS   Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
REFERENCE   5  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
  REMARK    Sequence update by submitter
COMMENT     On Apr 23, 2002 this sequence version replaced gi:6715112.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100490"
                     /tissue_type="venom"
     CDS             89. .547
                     /note="consists of two heterologous subunits (A and B
                     chains) linked by an inter-subunit disulfide bond; c-type
                     lectin family member"
                     /codon_start=1
                     /product="agkisacutacin A chain"
                     /protein_id="AAF26286.2"
                     /db_xref="GI:20273044"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCSSGWSSYEGHCYKVFKQSK
                     TWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHIETGFHKWENFYCEQQDPFVCEA"
BASE COUNT      191 a    173 c    204 g    165 t
ORIGIN

Alignment Scores:
Pred. No.:                1.45e-74       Length:          733
Score:                    701.00         Matches:         127
Percent Similarity:       98.45%         Conservative:    0
Best Local Similarity:    98.45%         Mismatches:      2
Query Match:              98.04%         Indels:          0
DB:                       5              Gaps:            0

US-09-938-114-2 (1-129) x AF176420 (1-733)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
                ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 GATTGTTCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAGGTCTTCAAACAATCT 217

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AAGACCTGGACAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 277

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGGCCAGTTGATTGCTCAGAAGATAAAG 337

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 TCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 397

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        398 ATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAA 457

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 AAGTGTCTTGGGGTGCACATAGAGACAGGGTTTCATAAGTGGGAGAATTTTTACTGTGAA 517

Qy        121 GlnGlnAspProPheValCysGluAla 129
              |||||||||||||||||||||||||||
Db        518 CAACAAGATCCTTTTGTCTGCGAGGCA 544

RESULT 2
AY091758
LOCUS       AY091758                497 bp    mRNA    linear   VRT 13-MAY-2002
DEFINITION  Deinagkistrodon acutus clone 2100590 ACF 1/2 A-chain mRNA, complete
            cds.
ACCESSION   AY091758
VERSION     AY091758.1  GI:20562936
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     A chain of ACF 1/2 from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 497)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100590"
     CDS             1. .459
                     /note="C-type lectin family member"
                     /codon_start=1
                     /product="ACF 1/2 A-chain"
                     /protein_id="AAM22786.1"
                     /db_xref="GI:20562937"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCSSSWSSYEGHCYKAFKQSK
                     TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
BASE COUNT      135 a    102 c    143 g    117 t
ORIGIN

Alignment Scores:
Pred. No.:                1.68e-69       Length:          497
Score:                    657.00         Matches:         119
Percent Similarity:       93.02%         Conservative:    1
Best Local Similarity:    92.25%         Mismatches:      9
```

```
Query Match:            91.89%      Indels:        0
DB:                     5           Gaps:          0

US-09-938-114-2 (1-129) x AY091758 (1-497)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||||||||   ||||||||||||||||||||||||||||||   ||||||||||||
Db         70 GATTGTTCCTCTAGTTGGTCCTCCTATGAAGGGCATTGCTACAAGGCCTTCAAACAATCT 129

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 AAGACCTGGGCAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 189

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db        190 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGCCCATCTGATTGCTCAGAAGATAAAG 249

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 TCAGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 309

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        310 ATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAA 369

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||||||||||||||      |||||||||   ||||||||||||||||||||||||
Db        370 AAGTGTCTTGGGGTGCACAAAGCGACAGGGTTTCGTAAGTGGGAGAATTTTTACTGTGAA 429

Qy        121 GlnGlnAspProPheValCysGluAla 129
              |||:::|||||||||||||||||||||
Db        430 CAACGAGATCCTTTCGTCTGCGAGGCA 456

RESULT 3
AB036880
LOCUS       AB036880                 656 bp    mRNA    linear   VRT 23-MAY-2002
DEFINITION  Deinagkistrodon acutus acp-a mRNA for anticoagulant protein A,
            complete cds.
ACCESSION   AB036880
VERSION     AB036880.1  GI:8980618
KEYWORDS    anticoagulant protein A.
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1
  AUTHORS   Tani,A., Ogawa,T., Nose,T., Nikandrov,N.N., Deshimaru,M.,
            Chijiwa,T., Chang,C.C., Fukumaki,Y. and Ohno,M.
  TITLE     Characterization, primary structure and molecular evolution of
            anticoagulant protein from Agkistrodon actus venom
  JOURNAL   Toxicon 40 (6), 803-813 (2002)
  MEDLINE   21975857
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Ogawa,T. and Tani,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate
            School of Agricultural Science; 1-1, Tsutusmidori Amamiyamachi,
            Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
            URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808,
            Fax:81-22-717-8807)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /tissue_type="venom gland"
     gene            1. .656
                     /gene="acp-a"
     CDS             50. .508
                     /gene="acp-a"
                     /codon_start=1
                     /product="anticoagulant protein A"
                     /protein_id="BAA99281.1"
                     /db_xref="GI:8980619"
                     /translation="MGRFIFVSFGLLVVYLSLSGTAADCSSSWSSYEGHCYKAFKQSK
                     TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
     sig_peptide     50. .118
                     /gene="acp-a"
BASE COUNT      173 a    148 c    185 g    150 t
ORIGIN

Alignment Scores:
Pred. No.:              2.36e-69        Length:         656
Score:                  657.00          Matches:        119
Percent Similarity:     93.02%          Conservative:   1
Best Local Similarity:  92.25%          Mismatches:     9
Query Match:            91.89%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-2 (1-129) x AB036880 (1-656)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||||||||   ||||||||||||||||||||||||||||||   ||||||||||||
Db        119 GATTGTTCCTCTAGTTGGTCCTCCTATGAAGGGCATTGCTATAAGGCCTTCAAACAATCT 178

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 AAGACCTGGGCAGATGCAGAGAGCTTCTGCACGAAGCAGGTGAACGGGGGGCATCTGGTC 238

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db        239 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGCCCATCTGATTGCTCAGAAGATAAAG 298

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TCGGCCAAAATCCATGTCTGGATCGGACTGAGGGCTCAAAACAAAGAAAAGCAATGCAGC 358

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        359 ATAGAGTGGAGCGATGGCTCCAGCATCAGTTATGAGAATTGGATTGAAGAAGAATCCAAA 418

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||||||||||||||      |||||||||   ||||||||||||||||||||||||
Db        419 AAGTGTCTTGGGGTGCACAAAGCGACAGGGTTTCGTAAGTGGGAGAATTTTTACTGTGAA 478

Qy        121 GlnGlnAspProPheValCysGluAla 129
              |||:::|||||||||||||||||||||
Db        479 CAACGAGATCCATTTGTATGCGAAGCA 505

RESULT 4
AR259035
LOCUS       AR259035                 544 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 1 from patent US 6489451.
ACCESSION   AR259035
VERSION     AR259035.1  GI:27309520
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 544)
  AUTHORS   Li,B.X. and Cheng,X.
  TITLE     Antithrombosis enzyme from the snake venom of agkistrodon acutus
  JOURNAL   Patent: US 6489451-A 1 03-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .544
                     /organism="unknown"
BASE COUNT      158 a    122 c    143 g    121 t
ORIGIN

Alignment Scores:
Pred. No.:              5.67e-69        Length:         544
Score:                  653.00          Matches:        119
Percent Similarity:     99.17%          Conservative:   0
Best Local Similarity:  99.17%          Mismatches:     1
```

```
Query Match:            91.33%      Indels:        0
DB:                     6           Gaps:          0

US-09-938-114-2 (1-129) x AR259035 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         30 CysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAsp 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy         50 PheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGly 69
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy         70 LeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIle 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy         90 SerLysGluAsnTrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThr 109
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        110 GlyPheHisLysTrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGluAla 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

RESULT 5
TFLFIXA
LOCUS       TFLFIXA                  678 bp    mRNA    linear   VRT 06-FEB-1999
DEFINITION  Trimeresurus flavoviridis mRNA for factor IX/factor X binding
            protein A chain, complete cds.
ACCESSION   D83331
VERSION     D83331.1  GI:1402639
KEYWORDS    facotr IX/factor X binding protein A chain.
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (bases 1 to 678)
  AUTHORS   Matsuzaki,R., Yoshiara,E., Yamada,M., Shima,K., Atoda,H. and
            Morita,T.
  TITLE     cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
            protein from snake venom
  JOURNAL   Biochem. Biophys. Res. Commun. 220 (2), 382-387 (1996)
  MEDLINE   96184662
   PUBMED   8645314
REFERENCE   2  (bases 1 to 678)
  AUTHORS   Morita,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1996) Takashi Morita, Meiji College of Pharmacy,
            Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
            (Tel:0424-21-0101(ex.429), Fax:0424-21-1489)
FEATURES             Location/Qualifiers
     source          1. .678
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
     5'UTR           1. .49
     CDS             50. .508
                     /codon_start=1
                     /product="factor IX/factor X binding protein A chain"
                     /protein_id="BAA11887.1"
                     /db_xref="GI:1402640"
                     /translation="MGRFIFMSFGLLVVAASLRGTGADCLSGWSSYEGHCYKAFEKYK
                     TWEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIWIGLRVQGKVKQC
                     NSEWSDGSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCGQQNPFVCEA"
     3'UTR           509. .678
     polyA_signal    659. .664
BASE COUNT      174 a    161 c    191 g    152 t

ORIGIN

Alignment Scores:
Pred. No.:              4.8e-50     Length:        678
Score:                  496.00      Matches:       87
Percent Similarity:     75.97%      Conservative:  11
Best Local Similarity:  67.44%      Mismatches:    31
Query Match:            69.37%      Indels:        0
DB:                     5           Gaps:          0

US-09-938-114-2 (1-129) x TFLFIXA (1-678)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||||||||||||||||||||   |||::::::
Db        119 GATTGTCTCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAGGCCTTCGAAAAATAC 178

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||  ||||||||     ||||||:::|||     |||  |||||||||
Db        179 AAGACCTGGGAAGATGCAGAGAGGGTCTGCACGGAGCAGGCGAAGGGCGCGCATCTGGTC 238

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||||||||||||||||   ||||||:::   |||   :::|||
Db        239 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGCCCAGCTGGTCACTCAGAACATGAAG 298

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                      ::::::|||||||||||||||||   |||   |||   |||||||||:::
Db        299 AGACTCGATTTCTATATCTGGATCGGGCTGAGGGTTCAAGGCAAAGTCAAGCAATGCAAC 358

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                 |||||||||||||||||:::|||   ||||||||||||||   |||||||||
Db        359 TCGGAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 418

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||||||||:::      ||||||   |||   ||||||   |||   ||||||
Db        419 ACGTGTCTTGGGCTGGAAAAAGAGACAGATTTTCGTAAGTGGGTCAATATTTACTGTGGA 478

Qy        121 GlnGlnAspProPheValCysGluAla 129
              ||||||:::||||||||||||||||||
Db        479 CAACAAAATCCTTTCGTCTGCGAGGCA 505

RESULT 6
AB046491
LOCUS       AB046491                 683 bp    mRNA    linear   VRT 27-JAN-2001
DEFINITION  Trimeresurus flavoviridis mRNA for factor XI/factor X binding
            protein A chain, complete cds.
ACCESSION   AB046491
VERSION     AB046491.1  GI:12583676
KEYWORDS    .
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (sites)
  AUTHORS   Tani,A., Ogawa,T., Nose,T., Nikandorov,N.N., Deshimaru,M.,
            Chijiwa,T., Chang,C., Fukumaki,Y. and Ohno,M.
  TITLE     Characterization and molecular evolution of an anticoagulant
            protein from Agkistrodon actus venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 683)
  AUTHORS   Ogawa,T. and Ohno,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-2000) Tomohisa Ogawa, Tohoku University, Dept. of
            Agricultural Science; 1-1 Tsutsumidori Amamiyamachi, Aoba-ku,
            Sendai, Miyagi 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
            Tel:81-22-717-8808, Fax:81-22-717-8807)
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
                     /tissue_type="venome gland"
                     /country="Japan:Kagoshima, Tokunoshima island"
```

```
     gene            1. .683
                     /gene="XI/Xbp-a"
     CDS             77. .517
                     /gene="XI/Xbp-a"
                     /codon_start=1
                     /product="factor XI/factor X binding protein A chain"
                     /protein_id="BAB21452.1"
                     /db_xref="GI:12583677"
                     /translation="MSFGLLVVAASLRGTGADCLSGWSSYEGHCYKAFEKYKTWEDAE
                     RVCTEQAKGAHLVSIESSGEADFVAQLVTQNMKRLDFYIWIGLRVQGKVKQCNSEWSD
                     GSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCGQQNPFVCEA"
BASE COUNT      172 a    162 c    194 g    155 t
ORIGIN

Alignment Scores:
Pred. No.:                4.85e-50      Length:          683
Score:                    496.00        Matches:         87
Percent Similarity:       75.97%        Conservative:    11
Best Local Similarity:    67.44%        Mismatches:      31
Query Match:              69.37%        Indels:          0
DB:                       5             Gaps:            0

US-09-938-114-2 (1-129) x AB046491 (1-683)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||  ||||||||||||||||||||||||||||||   |||:::::
Db        128 GATTGTCTCTCTGGTTGGTCCTCCTATGAGGGGCATTGCTACAAGGCCTTCGAAAAATAC 187

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||   ||||||||   ||||||:::|||   |||   |||||||||
Db        188 AAGACCTGGGAAGATGCAGAGAGGGTCTGCACGGAGCAGGCGAAGGGCGCGCATCTGGTC 247

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||||||||||||||||||||||||||||||   ||||||:::   |||   :::|||
Db        248 TCTATCGAAAGCTCCGGAGAAGCAGACTTTGTGGCCCAGCTGGTCACTCAGAACATGAAG 307

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                    ::::::|||||||||||||||   |||   |||   |||||||||:::
Db        308 AGACTCGATTTCTATATCTGGATCGGGCTGAGGGTTCAAGGAAAGGTCAAGCAATGCAAC 367

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                 ||||||||||||||||||||:::|||   ||||||||||||||   |||||||||
Db        368 TCGGAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 427

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
               |||||||||:::      ||||||   |||   ||||||   |||   ||||||
Db        428 ACGTGTCTTGGGCTGGAAAAAGAGACAGATTTTCGTAAGTGGGTCAATATTTACTGTGGA 487

Qy        121 GlnGlnAspProPheValCysGluAla 129
               ||||||:::||||||||||||||||||
Db        488 CAACAAAATCCTTTCGTCTGCGAGGCA 514

RESULT 7
AX427204
LOCUS       AX427204                721 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 4 from Patent WO0214364.
ACCESSION   AX427204
VERSION     AX427204.1  GI:21530566
KEYWORDS    .
SOURCE      Sistrurus miliarius (pigmy rattlesnake)
  ORGANISM  Sistrurus miliarius
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Sistrurus.
REFERENCE   1
  AUTHORS   Sheppard,P.O. and Bishop,P.D.
  TITLE     Rattlesnake venom gland proteins
  JOURNAL   Patent: WO 0214364-A 4 21-FEB-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .721
                     /organism="Sistrurus miliarius"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8758"
     CDS             91. .549
                     /note="Zsnk3"
                     /codon_start=1
                     /protein_id="CAD35499.1"
                     /db_xref="GI:21530567"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADCPSGWSSYDQHCYRVFKQLK
                     TWDDAERFCSEQAEGGHLVSIESSEEAAFVAQLVPENRRRAILYIWIGLRVQGKEKQC
                     SAKWSDGSSVSYENWIEAESKTCLGLQQGTNYHKWVNIYCGEINPFVCEA"
BASE COUNT      192 a    167 c    203 g    159 t
ORIGIN

Alignment Scores:
Pred. No.:                9.79e-48      Length:          721
Score:                    477.00        Matches:         83
Percent Similarity:       76.74%        Conservative:    16
Best Local Similarity:    64.34%        Mismatches:      30
Query Match:              66.71%        Indels:          0
DB:                       6             Gaps:            0

US-09-938-114-2 (1-129) x AX427204 (1-721)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||  ||||||||||||:::   ||||||||:::||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||   ||||||||   ||||||:::|||   ||||||||||||||
Db        220 AAGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTC 279

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||||||||||||||   ||||||   ||||||   ||||||:::   :::      :::
Db        280 TCTATCGAAAGCTCCGAAGAAGCAGCCTTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGG 339

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 |||   :::::::::|||||||||||||||   |||   ||||||||||||||||||
Db        340 AGAGCCATTCTCTATATCTGGATCGGACTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGC 399

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
               :::|||||||||||||||||:::|||   ||||||||||||||   |||||||||
Db        400 GCGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 459

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
               |||||||||:::          |||   :::|||||||||   |||   ||||||
Db        460 ACATGTCTTGGGCTGCAACAAGGCACAAATTATCATAAGTGGGTCAATATTTACTGTGGA 519

Qy        121 GlnGlnAspProPheValCysGluAla 129
               :::   :::||||||||||||||||||
Db        520 GAAATAAATCCTTTTGTCTGCGAGGCA 546

RESULT 8
AF190827
LOCUS       AF190827                630 bp    mRNA    linear   VRT 01-OCT-2000
DEFINITION  Gloydius halys halyxin A-chain precursor (HXNA) mRNA, complete cds.
ACCESSION   AF190827
VERSION     AF190827.1  GI:10441755
KEYWORDS    .
SOURCE      Gloydius halys (halys viper)
  ORGANISM  Gloydius halys
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Gloydius.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
  TITLE     A novel coagulation factor Xa inhibitor from Korean snake
            (Agkistrodon halys) venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 630)
  AUTHORS   Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-SEP-1999) Cardiovascular Center, Yonsei University
            College of Medicine, 134 Shinchon-dong, Seoul 120-752, Republic of
            Korea
```

```
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Gloydius halys"
                     /mol_type="mRNA"
                     /db_xref="taxon:8714"
     gene            1. .630
                     /gene="HXNA"
     CDS             24. .482
                     /gene="HXNA"
                     /function="coagulation factor Xa inhibitor"
                     /codon_start=1
                     /product="halyxin A-chain precursor"
                     /protein_id="AAG17178.1"
                     /db_xref="GI:10441756"
                     /translation="MGRFIFLSFGLLVVFLSLSGTGADCPSGWSSYEGHCYNIFHLFK
                     TWAEAERFCRKQVKGAHLVSIESSEEADFVAQLVSENMKRYGIYIWIGLRVRGKKKQC
                     SSQWSDGSSVSYQNWIEAESKTCLGLQKETEFRKWFNIYCGERNPFVCEA"
     sig_peptide     24. .92
                     /gene="HXNA"
     mat_peptide     93. .479
                     /gene="HXNA"
                     /product="halyxin A-chain"
BASE COUNT      173 a    137 c    174 g    146 t
ORIGIN

Alignment Scores:
Pred. No.:              1.44e-47        Length:         630
Score:                  475.00          Matches:        82
Percent Similarity:     76.74%          Conservative:   17
Best Local Similarity:  63.57%          Mismatches:     30
Query Match:            66.43%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-2 (1-129) x AF190827 (1-630)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||||||||||||||||||   :::|||
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAACATCTTCCATCTATTC 152

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||   :::||||||   ||||||   |||||||||   |||   |||||||||
Db        153 AAGACCTGGGCAGAAGCAGAGAGGTTCTGCAGGAAGCAGGTGAAGGGCGCGCATCTGGTC 212

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||   ||||||||||||||   ||||||:::::::::   :::|||
Db        213 TCTATCGAAAGCTCCGAAGAAGCAGACTTTGTGGCCCAGCTGGTCTCTGAGAACATGAAG 272

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 |||::::::|||||||||||||||   :::   |||:::|||||||||||||
Db        273 AGATACGGGATCTATATCTGGATCGGACTGAGGGTTCGAGGCAAAAAGAAGCAATGCAGC 332

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||:::|||   :::|||||||||||||   |||||||||
Db        333 TCCCAGTGGAGCGATGGCTCCAGCGTCAGTTATCAGAACTGGATTGAAGCAGAATCCAAA 392

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||||||||:::      ||||||   |||   ||||||   |||   ||||||
Db        393 ACGTGTCTTGGGCTGCAAAAAGAGACAGAGTTTCGTAAGTGGTTTAATATTTACTGTGGA 452

Qy        121 GlnGlnAspProPheValCysGluAla 129
              :::::::::||||||||||||||||||
Db        453 GAACGAAATCCTTTCGTCTGCGAGGCA 479

RESULT 9
AF125309
LOCUS       AF125309                632 bp    mRNA    linear   VRT 04-MAR-1999
DEFINITION  Gloydius halys brevicaudus fibrinogen clotting inhibitor A chain
            mRNA, complete cds.
ACCESSION   AF125309
VERSION     AF125309.1  GI:4337049
KEYWORDS    .
SOURCE      Gloydius halys brevicaudus
  ORGANISM  Gloydius halys brevicaudus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Gloydius.
REFERENCE   1  (bases 1 to 632)
  AUTHORS   Kim,D.S. and Koh,Y.S.
  TITLE     Purification and molecular cloning of snake venom fibrin clotting
            inhibitor
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 632)
  AUTHORS   Kim,D.S. and Koh,Y.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-1999) Biochemistry, Yonsei University, Seodaemun
            ku Shinchon dong 134, Seoul 120-749, Korea
FEATURES             Location/Qualifiers
     source          1. .632
                     /organism="Gloydius halys brevicaudus"
                     /mol_type="mRNA"
                     /sub_species="brevicaudus"
                     /db_xref="taxon:66175"
                     /note="venom"
     CDS             34. .498
                     /codon_start=1
                     /product="fibrinogen clotting inhibitor A chain"
                     /protein_id="AAD18055.1"
                     /db_xref="GI:4337050"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADFFCPSGWGSNNGHCYQAFNQ
                     RMTWEDAERFCSAQAKGGHLVSIETRAEADFVAHVVAERIETSFPHVWIGLRDEGKEQ
                     QCSSEWSDGSSVSYENWIEAESKTCLGLELDSNYHKWVNVYCGQRNPFVCEA"
BASE COUNT      162 a    140 c    182 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              9.07e-46        Length:         632
Score:                  460.00          Matches:        77
Percent Similarity:     75.78%          Conservative:   20
Best Local Similarity:  60.16%          Mismatches:     31
Query Match:            64.34%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-2 (1-129) x AF125309 (1-632)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||   |||   |||   |||       ||||||||||:::   |||   |||
Db        112 TGTCCCTCTGGTTGGGGCAGCAATAATGGTCATTGCTACCAGGCCTTCAATCAACGGATG 171

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              ||||||   |||||||||   ||||||:::   |||       ||||||||||||||||||
Db        172 ACCTGGGAAGATGCAGAGAGGTTCTGCTCAGCACAGGCGAAGGGCGGGCATCTGGTCTCT 231

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              ||||||:::      ||||||||||||||       :::::::|||:::::::|||::::::
Db        232 ATCGAAACACGGGCAGAAGCAGACTTTGTGGCCCACGTGGTCGCTGAGAGGATAGAGACA 291

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
              :::      ||||||||||||||||||||   :::   ||||||:::||||||||||
Db        292 TCCTTTCCCCATGTCTGGATCGGACTGAGGGATGAAGGCAAAGAACAGCAATGCAGCTCT 351

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLysLys 101
              ||||||||||||||||||||:::|||   ||||||||||||||   |||||||||
Db        352 GAGTGGAGCGATGGATCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAAACA 411

Qy        102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
              ||||||||:::   :::::::::   :::||||||||   |||   ||||||   |||
Db        412 TGTCTTGGGCTGGAACTAGACTCAAATTATCATAAGTGGGTCAATGTTTACTGTGGACAA 471

Qy        122 GlnAspProPheValCysGluAla 129
              ::::::||||||||||||||||||
Db        472 AGAAATCCTTTCGTCTGCGAGGCT 495

RESULT 10
AF463522
LOCUS       AF463522                508 bp    mRNA    linear   VRT 21-JAN-2002
DEFINITION  Deinagkistrodon acutus antithrombin 1 A chain mRNA, complete cds.
```

```
ACCESSION   AF463522
VERSION     AF463522.1  GI:18252679
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Yu,H., Xiang,K. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-DEC-2001) Dept. of Biochemistry and Molecular
            Biology, School of Life-Science, Huangshan Road, Hefei, Anhui
            230027, China
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="3015"
     CDS             1. .477
                     /note="C-type lectin family member"
                     /codon_start=1
                     /product="antithrombin 1 A chain"
                     /protein_id="AAL66391.1"
                     /db_xref="GI:18252680"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADFNCPPGWSAYDQYCYQVIKE
                     PKNWDDAERFCTEQADGGHLVSIESKGERDFVAQLVSQSIESVEDHVWTGLRVQNKEK
                     QCSTEWSDGSSVSYENLLELYMRKCGALERETGFHKWINLGCIQLNPFVCKFPPQC"
BASE COUNT      131 a    107 c    151 g    119 t
ORIGIN
```

Alignment Scores:

Pred. No.:	2.28e-43	Length:	508
Score:	439.00	Matches:	77
Percent Similarity:	74.22%	Conservative:	18
Best Local Similarity:	60.16%	Mismatches:	33
Query Match:	61.40%	Indels:	0
DB:	5	Gaps:	0

US-09-938-114-2 (1-129) x AF463522 (1-508)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          ||||||:::|||:::   :::||||||:::|||   |||:::
Db         76 AATTGTCCCCCTGGTTGGTCCGCCTATGATCAGTATTGCTACCAGGTCATCAAAGAACCG 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||   |||   |||||||||   |||||||||:::|||   :::|||||||||||||||
Db        136 AAAAACTGGGATGATGCAGAGCGGTTCTGCACGGAGCAGGCGGACGGTGGGCATCTGGTC 195

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||   ||||||   |||||||||   ||||||:::::::|||   |||:::
Db        196 TCTATCGAAAGCAAGGGAGAGAGAGACTTTGTGGCCCAGCTGGTCTCTCAGAGCATAGAG 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||   :::   |||||||||   |||||||||   ||||||||||||||||||||||||
Db        256 AGTGTTGAAGACCATGTCTGGACCGGACTGAGGGTTCAAAACAAAGAAAAGCAATGCAGC 315

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              ||||||||||||||||||||||:::|||   ||||||   :::|||          :::
Db        316 ACAGAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACCTGCTTGAACTATATATGAGA 375

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||    :::    ||||||||||||||||||||   |||    |||
Db        376 AAGTGTGGGGCGCTGGAACGAGAGACAGGGTTTCATAAGTGGATCAATCTTGGCTGTATA 435

Qy        121 GlnGlnAspProPheValCysGlu 128
              |||   :::|||||||||||:::
Db        436 CAACTGAATCCTTTCGTCTGCAAG 459
```

```
RESULT 11
AF540646
LOCUS       AF540646                661 bp    mRNA    linear   VRT 26-SEP-2002
DEFINITION  Deinagkistrodon acutus agglucetin-alpha 2 subunit precursor, mRNA,
            complete cds.
ACCESSION   AF540646
VERSION     AF540646.1  GI:23321260
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Wang,W.J. and Huang,T.F.
  TITLE     A novel tetrameric venom protein, agglucetin from Agkistrodon
            acutus, acts as a glycoprotein Ib agonist
  JOURNAL   Thromb. Haemost. 86 (4), 1077-1086 (2001)
  MEDLINE   21542539
   PUBMED   11686327
REFERENCE   2  (bases 1 to 661)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Molecular structure and functional characterization of agglucetin,
            a tetrameric glycoprotein Ib-binding protein, from Formosan pit
            viper
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 661)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
            of Technology, 261, Wen-Hwa 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
            333-03, R.O.C.
FEATURES             Location/Qualifiers
     source          1. .661
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
     CDS             13. .489
                     /note="platelet agglutination inducer; glycoprotein Ib
                     agonist"
                     /codon_start=1
                     /product="agglucetin-alpha 2 subunit precursor"
                     /protein_id="AAN23125.1"
                     /db_xref="GI:23321261"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADFNCPPGWSAYDQYCYQVIKE
                     PKNWDDAERFCTEQADGGHLVSIESKGERDFVAQLVSQNIESVEDHVWTGLRVQNKEK
                     QCSTEWSDGSSVSYENLLELYMRKCGALERETGFHKWINLGCIQLNPFVCKFPPQC"
     sig_peptide     13. .81
     mat_peptide     82. .486
                     /product="agglucetin-alpha 2 subunit"
BASE COUNT      167 a    159 c    183 g    152 t
ORIGIN
```

Alignment Scores:

Pred. No.:	3.15e-43	Length:	661
Score:	439.00	Matches:	77
Percent Similarity:	74.22%	Conservative:	18
Best Local Similarity:	60.16%	Mismatches:	33
Query Match:	61.40%	Indels:	0
DB:	5	Gaps:	0

US-09-938-114-2 (1-129) x AF540646 (1-661)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          ||||||:::|||:::   :::||||||:::|||   |||:::
Db         88 AATTGTCCCCCTGGTTGGTCCGCCTATGATCAGTATTGCTACCAGGTCATCAAAGAACCG 147

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||   |||   |||||||||   |||||||||:::|||   :::|||||||||||||||
Db        148 AAAAACTGGGATGATGCAGAGCGGTTCTGCACGGAGCAGGCGGACGGTGGGCACCTGGTC 207

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||   ||||||   |||||||||   ||||||:::::::|||   |||:::
Db        208 TCTATCGAAAGCAAGGGAGAGAGAGACTTTGTGGCCCAGCTGGTCTCTCAGAACATAGAG 267

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
```

```
Db        268 AGTGTTGAAGACCATGTCTGGACCGGACTGAGGGTTCAAAACAAAGAAAAGCAATGCAGC 327

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        328 ACAGAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACCTGCTTGAACTATATATGAGA 387

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        388 AAGTGTGGGGCGCTGGAACGAGAGACAGGGTTTCATAAGTGGATCAATCTTGGCTGTATA 447

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        448 CAACTGAATCCTTTCGTCTGCAAG 471

RESULT 12
AX427207
LOCUS       AX427207                 580 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 7 from Patent WO0214364.
ACCESSION   AX427207
VERSION     AX427207.1  GI:21530569
KEYWORDS    .
SOURCE      Sistrurus miliarius (pigmy rattlesnake)
  ORGANISM  Sistrurus miliarius
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Sistrurus.
REFERENCE   1
  AUTHORS   Sheppard,P.O. and Bishop,P.D.
  TITLE     Rattlesnake venom gland proteins
  JOURNAL   Patent: WO 0214364-A 7 21-FEB-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Sistrurus miliarius"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8758"
     CDS             <3. .437
                     /note="Zsnk4"
                     /codon_start=1
                     /protein_id="CAD35500.1"
                     /db_xref="GI:21530570"
                     /translation="IRNEGGTGADFDCPSDWYAYDQYCYRVIKQLRTWDDAERFCSEQ
                     AKGGHLVSIESDGEAAFVAQLVAENIKQNKYDVWIGLRIQGEEKQCSTKWSDGSSVNY
                     ENLIKHATKKCFGLKKETGFRTWRNVHCTQQNLFMCKFPPEC"
BASE COUNT      154 a    133 c    163 g    130 t
ORIGIN

Alignment Scores:
Pred. No.:                1.07e-42        Length:          580
Score:                    434.00          Matches:         75
Percent Similarity:       74.22%          Conservative:    20
Best Local Similarity:    58.59%          Mismatches:      33
Query Match:              60.70%          Indels:          0
DB:                       6               Gaps:            0

US-09-938-114-2 (1-129) x AX427207 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db         96 AGGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTC 155

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        156 TCTATTGAAAGCGACGGAGAAGCAGCCTTTGTGGCCCAGCTGGTCGCTGAGAACATCAAG 215

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        216 CAAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGGCGAAGAGAAGCAATGCAGC 275

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        276 ACGAAGTGGAGCGATGGCTCCAGCGTCAATTATGAGAACCTGATTAAACATGCGACCAAA 335

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        336 AAGTGTTTTGGGCTGAAAAAAGAGACAGGGTTTCGCACGTGGCGCAATGTTCACTGTACA 395

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        396 CAACAAAATCTTTTCATGTGCAAG 419

RESULT 13
AF244900
LOCUS       AF244900                 408 bp    mRNA    linear   VRT 29-JUN-2000
DEFINITION  Calloselasma rhodostoma aggretin alpha chain mRNA, partial cds.
ACCESSION   AF244900
VERSION     AF244900.1  GI:8809811
KEYWORDS    .
SOURCE      Calloselasma rhodostoma (Malayan pit viper)
  ORGANISM  Calloselasma rhodostoma
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Calloselasma.
REFERENCE   1  (bases 1 to 408)
  AUTHORS   Chung,C.H., Au,L.C. and Huang,T.F.
  TITLE     Molecular cloning and sequence analysis of aggretin, a
            collagen-like platelet aggregation inducer
  JOURNAL   Biochem. Biophys. Res. Commun. 263 (3), 723-727 (1999)
  MEDLINE   99443731
   PUBMED   10512747
REFERENCE   2  (bases 1 to 408)
  AUTHORS   Chung,C.H., Au,L.C. and Huang,T.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2000) Pharmacology, College of Medicine, National
            Taiwan University, No. 1, Sec. 1, Jen-Ai Rd., Taipei, Taiwan
FEATURES             Location/Qualifiers
     source          1. .408
                     /organism="Calloselasma rhodostoma"
                     /mol_type="mRNA"
                     /db_xref="taxon:8717"
                     /tissue_type="venom gland"
     CDS             <1. .>408
                     /note="forms heterodimer of alpha and beta chains;
                     collagen-receptor agonist; induces platlet aggregation;
                     similar to C-type lectins and glycoprotein Ib binding
                     proteins"
                     /codon_start=1
                     /product="aggretin alpha chain"
                     /protein_id="AAF79952.1"
                     /db_xref="GI:8809812"
                     /translation="GLEDCDFGWSPYDQHCYQAFNEQKTWDEAEKFCRAQENGAHLAS
                     IESNGEADFVSWLISQKDELADEDYVWIGLRAQNKEQQCSSEWSDGSSVSYENLIDLH
                     TKKCGALEKLTGFRKWVNYYCEQMHAFVCKLLPY"
BASE COUNT      115 a     79 c    115 g     99 t
ORIGIN

Alignment Scores:
Pred. No.:                2.87e-41        Length:          408
Score:                    420.50          Matches:         77
Percent Similarity:       72.09%          Conservative:    16
Best Local Similarity:    59.69%          Mismatches:      35
Query Match:              58.81%          Indels:          1
DB:                       5               Gaps:            1

US-09-938-114-2 (1-129) x AF244900 (1-408)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         10 GATTGTGACTTTGGTTGGTCTCCCTATGATCAGCATTGCTACCAGGCATTCAATGAACAA 69

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
```

```
Db         70 AAAACCTGGGATGAGGCAGAGAAGTTCTGCAGGGCGCAGGAGAATGGTGCGCATCTGGCC 129

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        130 TCTATCGAAAGCAATGGAGAAGCAGACTTTGTTTCCTGGCTGATTTCTCAGAAAGACGAA 189

Qy         61 SerAla---LysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCys 79

Db        190 CTGGCAGACGAAGACTACGTCTGGATCGGACTGAGGGCTCAAAACAAAGAACAGCAATGC 249

Qy         80 SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSer 99

Db        250 AGCTCGGAGTGGAGCGATGGCTCCAGTGTCAGTTATGAGAACTTGATTGATCTACATACG 309

Qy        100 LysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCys 119

Db        310 AAAAAGTGTGGTGCGCTGGAAAAACTGACAGGGTTTCGTAAGTGGGTCAATTATTACTGT 369

Qy        120 GluGlnGlnAspProPheValCysGlu 128

Db        370 GAACAAATGCATGCTTTCGTCTGCAAG 396

RESULT 14
AX427206
LOCUS       AX427206                456 bp    DNA     linear   PAT 19-JUN-2002
DEFINITION  Sequence 6 from Patent WO0214364.
ACCESSION   AX427206
VERSION     AX427206.1  GI:21530568
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Sheppard,P.O. and Bishop,P.D.
  TITLE     Rattlesnake venom gland proteins
  JOURNAL   Patent: WO 0214364-A 6 21-FEB-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="This degenerate nucleotide sequence encodes the
                     amino acid sequence of SEQ ID NO:5."
BASE COUNT       72 a     31 c     92 g     66 t    195 others
ORIGIN

Alignment Scores:
Pred. No.:              2.15e-38        Length:         456
Score:                  397.00          Matches:        66
Percent Similarity:     59.69%          Conservative:   11
Best Local Similarity:  51.16%          Mismatches:     52
Query Match:            55.52%          Indels:         0
DB:                     6               Gaps:           0

US-09-938-114-2 (1-129) x AX427206 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        130 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNGARGGNGGNCAYYTNGTN 189

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        190 WSNATHGARWSNWSNGARGARGCNGCNTTYGTNGCNCARYTNGTNCCNGARAAYMGNMGN 249

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        250 MGNGCNATHYTNTAYATHTGGATHGGNYTNMGNGTNCARGGNAARGARAARCARTGYWSN 309

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        310 GCNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYTGGATHGARGCNGARWSNAAR 369

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        370 ACNTGYYTNGGNYTNCARCARGGNACNAAYTAYCAYAARTGGGTNAAYATHTAYTGYGGN 429

Qy        121 GlnGlnAspProPheValCysGluAla 129

Db        430 GARATHAAYCCNTTYGTNTGYGARGCN 456

RESULT 15
AY099321
LOCUS       AY099321                477 bp    mRNA    linear   VRT 29-MAY-2002
DEFINITION  Protobothrops mucrosquamatus C-type lectin-like protein TMVA A
            chain mRNA, complete cds.
ACCESSION   AY099321
VERSION     AY099321.1  GI:21260581
KEYWORDS    .
SOURCE      Protobothrops mucrosquamatus
  ORGANISM  Protobothrops mucrosquamatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Protobothrops.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Wei,Q., Lu,Q.M., Jin,Y., Li,R., Wei,J., Wang,W. and Xiong,Y.
  TITLE     Purification and cloning of a novel C-type lectin-like protein with
            platelet aggregation activity from Trimeresurus mucrosquamatus
            venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 477)
  AUTHORS   Wei,Q., Lu,Q.M., Jin,Y., Li,R., Wei,J., Wang,W. and Xiong,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Department of Animal Toxicology, Kunming
            Institute of Zoology, 32 East Jiaochang Road, Kunming, Yunnan
            650223, China
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Protobothrops mucrosquamatus"
                     /mol_type="mRNA"
                     /isolation_source="venom"
                     /db_xref="taxon:103944"
     CDS             1. .477
                     /note="platelet aggregation activator"
                     /codon_start=1
                     /product="C-type lectin-like protein TMVA A chain"
                     /protein_id="AAM43808.1"
                     /db_xref="GI:21260582"
                     /translation="MGRFTFVSFGLLVVFLSLSGTGADFDCIPGWSAYDRYCYQAFSE
                     PKNWEDAESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNKEQ
                     QCRSEWSDASSVNYENLFKQSSKKCYALKKGTELRTWFNVYCGRENPFVCKYTPEC"
BASE COUNT      122 a    100 c    134 g    121 t
ORIGIN

Alignment Scores:
Pred. No.:              5.2e-38         Length:         477
Score:                  394.00          Matches:        69
Percent Similarity:     70.31%          Conservative:   21
Best Local Similarity:  53.91%          Mismatches:     38
Query Match:            55.10%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-2 (1-129) x AY099321 (1-477)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         76 GATTGTATCCCTGGTTGGTCCGCCTATGATCGGTATTGCTACCAGGCCTTCAGTGAACCG 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        136 AAAAACTGGGAAGATGCAGAGAGTTTCTGCGAGGAGGGGGTGAAGACCTCGCATCTGGTC 195
```

```
Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
Db        196 TCCATCGAAAGCTCCGGAGAAGGAGACTTCGTGGCCCAGCTGGTCGCTGAGAAGATAAAG 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
Db        256 ACGTCCTTTCAGTATGTCTGGATTGGGCTGAGGATTCAAAACAAAGAACAGCAATGCAGG 315

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
Db        316 TCGGAGTGGAGCGATGCCTCCAGTGTCAATTATGAGAACTTGTTTAAACAATCTTCCAAA 375

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
Db        376 AAATGTTATGCGCTGAAAAAAGGGACAGAGCTTCGCACGTGGTTCAATGTTTACTGTGGA 435

Qy        121 GlnGlnAspProPheValCysGlu 128
Db        436 AGAGAAAATCCTTTCGTCTGCAAG 459
```

Search completed: December 8, 2003, 16:08:44
Job time : 5477.5 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:39:59 ; Search time 443.023 Seconds
(without alignments)
786.026 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22839/app_query.fasta_1.725
-DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_0_@runat_08122003_091000_22839 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	653	91.3	544	25	ABX93668	cDNA encoding Dein
	2	477	66.7	721	24	AAD32055	Pigmy rattlesnake
	3	475	66.4	601	22	AAI71876	Snake venom blood
	4	460	64.3	632	24	AAL42015	Korean adder snake
	5	434	60.7	580	24	AAD32057	Pigmy rattlesnake
	6	397	55.5	456	24	AAD32056	Pigmy rattlesnake
	7	387	54.1	725	24	AAD32059	Pigmy rattlesnake
	8	378	52.9	432	24	AAD32058	Pigmy rattlesnake
	9	329	46.0	474	24	AAD32060	Pigmy rattlesnake
	10	290	40.6	633	22	AAI71877	Snake venom blood
	11	286.5	40.1	690	16	AAQ89309	Snake venom antith
	12	286.5	40.1	690	21	AAC61144	DNA encoding a sna
	13	286.5	40.1	690	24	AAK99834	DNA encoding the a
	14	272.5	38.1	583	24	AAL42016	Korean adder snake
	15	264	36.9	483	24	AAD32053	Pigmy rattlesnake
	16	237.5	33.2	716	17	AAT64829	Snake venom protei
	17	216	30.2	454	25	ABX93674	cDNA encoding Dein
	18	210	29.4	453	24	AAD32054	Pigmy rattlesnake
	19	201.5	28.2	441	11	AAQ05622	Sequence encoding
	20	201.5	28.2	562	22	AAS31360	Human cDNA encodin
	21	201.5	28.2	562	24	ABQ66684	Human polynucleoti
	22	201.5	28.2	777	21	AAZ99820	cDNA sequence of h
	23	201.5	28.2	800	22	AAF75077	Human colon associ
	24	201.5	28.2	843	21	AAC77946	Human cancer assoc
	25	201.5	28.2	843	22	AAF75090	Human colon associ
	26	201.5	28.2	1066	22	AAS31201	Human cDNA encodin
	27	201.5	28.2	1066	24	ABQ66525	Human polynucleoti
	28	201.5	28.2	4771	18	AAV02186	Human type C lecti
	29	201.5	28.2	4771	21	AAA92447	Human type C lecti
	30	201.5	28.2	5641	24	ABL92107	Human Tumour Endot
	31	201.5	28.2	5641	25	ABX72032	DNA encoding human
	32	201.5	28.2	5641	25	ABZ34758	Coding sequence SE
	33	201.5	28.2	5747	21	AAC77360	Human ORFX ORF2915
	34	198.5	27.8	678	24	ABZ11665	Human polynucleoti
	35	197.5	27.6	498	9	AAN81962	Sequence of human
	36	197.5	27.6	4451	25	ABZ34683	Coding sequence SE
	37	197.5	27.6	4588	18	AAV02185	Murine type C lect
	38	197.5	27.6	4588	21	AAA92446	Murine E-selectin
	39	197.5	27.6	4588	24	ABL92140	Mouse Tumour Endot
	40	197.5	27.6	4588	25	ABX72065	DNA encoding mouse
	41	197.5	27.6	5837	21	AAA92454	Murine E-selectin
	42	196.5	27.5	495	14	AAQ40477	MUREG-1 coding reg
	43	195.5	27.3	495	10	AAN91109	Human reg cDNA. H
c	44	193.5	27.1	710	24	ABV95090	Human pancreatic c
c	45	192.5	26.9	539	24	ABV98260	Human pancreatic c

ALIGNMENTS

RESULT 1
ABX93668
ID ABX93668 standard; cDNA; 544 BP.
XX
AC ABX93668;
XX
DT 02-JUN-2003 (first entry)
XX
DE cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
KW Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
KW platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW peripheral arterial occulsion; stroke; atherosclerosis; inflammation;

```
KW   thrombosis.
XX
OS   Deinagkistrodon acutus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..363
FT                   /*tag= a
FT                   /product= "Antithrombosis enzyme alpha chain"
FT                   /partial
FT                   /transl_except= (pos:244..246,aa:Lys)
FT                   /note= "No start codon given"
XX
PN   US6489451-B1.
XX
PD   03-DEC-2002.
XX
PF   10-APR-1998;   98US-0058740.
XX
PR   10-APR-1997;   97US-043886P.
XX
PA   (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI   Li BX,  Cheng X;
XX
DR   WPI; 2003-352116/33.
DR   P-PSDB; ABU08798.
XX
PT   New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT   preventing and treating vaso-occulusive and thromboembolic disorders,
PT   including myocardial infarction, restenosis, cerebral thrombosis and
PT   unstable angina -
XX
PS   Example 3; Column 25-26; 19pp; English.
XX
CC   The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC   enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC   prevents platelet aggregation. The anti-thrombosis enzyme was
CC   administered to rabbits intravenously. Thrombosis was determined before
CC   and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC   Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC   following administration at 0.005 micro/kg and this activity was
CC   increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC   composition is useful for preventing and treating vaso-occulusive and
CC   thromboembolic disorders, including myocardial infarction, restenosis,
CC   angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC   diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC   thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC   peripheral arterial occulsion, stroke. It is also useful for treating
CC   atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC   diseases. The present sequence represents cDNA encoding the
CC   Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
SQ   Sequence 544 BP; 158 A; 122 C; 143 G; 121 T; 0 other;

Alignment Scores:
Pred. No.:              5.36e-69        Length:         544
Score:                  653.00          Matches:        119
Percent Similarity:     99.17%          Conservative:   0
Best Local Similarity:  99.17%          Mismatches:     1
Query Match:            91.33%          Indels:         0
DB:                     25              Gaps:           0

US-09-938-114-2 (1-129) x ABX93668 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         30 CysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAsp 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy         50 PheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGly 69
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy         70 LeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIle 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy         90 SerLysGluAsnTrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThr 109
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        110 GlyPheHisLysTrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGluAla 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

RESULT 2
AAD32055
ID   AAD32055 standard; DNA; 721 BP.
XX
AC   AAD32055;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk3 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk3; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             91..549
FT                   /*tag= a
FT                   /product= "Zsnk3 protein"
FT   sig_peptide     91..159
FT                   /*tag= b
FT   mat_peptide     160..546
FT                   /*tag= c
FT                   /product= "Mature Zsnk3 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20179.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology -
XX
PS   Claim 5; Page 73-74; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
```

```
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk3 gene.
XX
SQ   Sequence 721 BP; 192 A; 167 C; 203 G; 159 T; 0 other;

Alignment Scores:
Pred. No.:                1.04e-47      Length:          721
Score:                    477.00        Matches:         83
Percent Similarity:       76.74%        Conservative:    16
Best Local Similarity:    64.34%        Mismatches:      30
Query Match:              66.71%        Indels:          0
DB:                       24            Gaps:            0

US-09-938-114-2 (1-129) x AAD32055 (1-721)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||:::   |||||||||:::|||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   |||||||||   ||||||::::::|||       ||||||||||||||
Db        220 AAGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTC 279

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||   ||||||   ||||||   |||||:::   :::      :::
Db        280 TCTATCGAAAGCTCCGAAGAAGCAGCCTTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGG 339

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||   :::::::::|||||||||||||||   |||   ||||||||||||||||||
Db        340 AGAGCCATTCTCTATATCTGGATCGGACTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGC 399

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||||:::|||   ||||||||||||||   ||||||||
Db        400 GCGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 459

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||||||||:::          |||   :::|||||||||   |||   ||||||
Db        460 ACATGTCTTGGGCTGCAACAAGGCACAAATTATCATAAGTGGGTCAATATTTACTGTGGA 519

Qy        121 GlnGlnAspProPheValCysGluAla 129
              :::   :::||||||||||||||||||
Db        520 GAAATAAATCCTTTTGTCTGCGAGGCA 546

RESULT 3
AAI71876
ID   AAI71876 standard; DNA; 601 BP.
XX
AC   AAI71876;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin A chain coding sequence.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis; ds.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   P-PSDB; AAM51543.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 1; Page 9; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   encodes the A chain of halyxin.
XX
SQ   Sequence 601 BP; 166 A; 133 C; 163 G; 139 T; 0 other;

Alignment Scores:
Pred. No.:                1.41e-47      Length:          601
Score:                    475.00        Matches:         82
Percent Similarity:       76.74%        Conservative:    17
Best Local Similarity:    63.57%        Mismatches:      30
Query Match:              66.43%        Indels:          0
DB:                       22            Gaps:            0

US-09-938-114-2 (1-129) x AAI71876 (1-601)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   |||||||||||||||||||||||||||||   :::|||
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAACATCTTCCATCTATTC 152

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   :::|||||||   ||||||   |||||||||   |||   |||||||||
Db        153 AAGACCTGGGCAGAAGCAGAGAGGTTCTGCAGGAAGCAGGTGAAGGGCGCGCATCTGGTC 212

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||||   ||||||||||||||||   ||||||:::::::::   :::|||
Db        213 TCTATCGAAAGCTCCGAAGAAGCAGACTTTGTGGCCCAGCTGGTCTCTGAGAACATGAAG 272

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 |||::::::|||||||||||||||||   :::   |||:::||||||||||||
Db        273 AGATACGGGATCTATATCTGGATCGGACTGAGGGTTCGAGGCAAAAAGAAGCAATGCAGC 332

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||||:::|||   :::||||||||||||   |||||||||
Db        333 TCCCAGTGGAGCGATGGCTCCAGCGTCAGTTATCAGAACTGGATTGAAGCAGAATCCAAA 392

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||||||||:::       ||||||   |||   ||||||   |||   ||||||
Db        393 ACGTGTCTTGGGCTGCAAAAAGAGACAGAGTTTCGTAAGTGGTTTAATATTTACTGTGGA 452

Qy        121 GlnGlnAspProPheValCysGluAla 129
              :::::::::||||||||||||||||||
Db        453 GAACGAAATCCTTTCGTCTGCGAGGCA 479

RESULT 4
AAL42015
ID   AAL42015 standard; cDNA; 632 BP.
XX
AC   AAL42015;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Korean adder snake venom salmorin A chain protein cDNA sequence.
XX
KW   Korean adder; ss; salmorin protein A chain; snake venom;
KW   fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW   thrombin binding; blood coagulation.
XX
OS   Agkistrodon halys brevicaudus.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             34..498
FT                   /*tag=  a
FT                   /product= "Salmorin A chain protein"
FT   sig_peptide     34..102
FT                   /*tag=  b
FT                   /note= "Signal peptide"
FT   mat_peptide     103..495
FT                   /*tag=  c
FT                   /note= "Mature Salmorin A chain protein"
FT   3'UTR           499..622
FT                   /*tag=  d
FT   polyA_signal    623..632
FT                   /*tag=  e
XX
PN   WO200214514-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-KR01277.
XX
PR   27-JUL-2000; 2000KR-0043470.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Chung K,  Kim D,  Koh Y;
XX
DR   WPI; 2002-241907/29.
DR   P-PSDB; AAO14520.
XX
PT   New salmorin protein derived from venom of Korean adder Agkistrodon
PT   halys brevicaudus, useful for treating thrombosis by repressing
PT   fibrinogen clotting through repression of activation of prothrombin
PT   into thrombin  -
XX
PS   Claim 1; Fig 1A; 30pp; English.
XX
CC   The invention comprises the nucleotide and protein sequences of a
CC   salmorin protein derived from the venom of Korean adder. Salmorin protein
CC   is composed of an A chain and a B chain, and has inhibitory activity
CC   against fibrinogen clotting. The salmorin protein of the invention is
CC   useful for treating thrombosis, as it represses fibrinogen clotting
CC   potently by binding to prothrombin and thrombin so as to delay blood
CC   coagulation. The present cDNA sequence encodes the Korean adder salmorin
CC   A chain protein.
XX
SQ   Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;

Alignment Scores:
Pred. No.:                9.53e-46        Length:          632
Score:                    460.00          Matches:         77
Percent Similarity:       75.78%          Conservative:    20
Best Local Similarity:    60.16%          Mismatches:      31
Query Match:              64.34%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-2 (1-129) x AAL42015 (1-632)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||   |||   |||   |||     ||||||||||:::   |||   |||
Db        112 TGTCCCTCTGGTTGGGGCAGCAATAATGGTCATTGCTACCAGGCCTTCAATCAACGGATG 171

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              ||||||  |||||||||  ||||||:::  |||      ||||||||||||||||||
Db        172 ACCTGGGAAGATGCAGAGAGGTTCTGCTCAGCACAGGCGAAGGGCGGGCATCTGGTCTCT 231

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              ||||||:::      ||||||||||||||      :::::::|||:::::::|||::::::
Db        232 ATCGAAACACGGGCAGAAGCAGACTTTGTGGCCCACGTGGTCGCTGAGAGGATAGAGACA 291

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
              :::      ||||||||||||||||||   :::   ||||||:::||||||||||
Db        292 TCCTTTCCCCATGTCTGGATCGGACTGAGGGATGAAGGCAAAGAACAGCAATGCAGCTCT 351

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLysLys 101
              ||||||||||||||||||||:::|||   ||||||||||||||   ||||||||
Db        352 GAGTGGAGCGATGGATCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAAACA 411

Qy        102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
              |||||||||:::   :::::::::   :::||||||||||   |||   ||||||   |||
Db        412 TGTCTTGGGCTGGAACTAGACTCAAATTATCATAAGTGGGTCAATGTTTACTGTGGACAA 471

Qy        122 GlnAspProPheValCysGluAla 129
              ::::::||||||||||||||||||
Db        472 AGAAATCCTTTCGTCTGCGAGGCT 495

RESULT 5
AAD32057
ID   AAD32057 standard; DNA; 580 BP.
XX
AC   AAD32057;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk4 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk4; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..437
FT                   /*tag=  a
FT                   /product= "Zsnk4 protein"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   sig_peptide     3..29
FT                   /*tag=  b
FT   mat_peptide     30..434
FT                   /*tag=  c
FT                   /product= "Mature Zsnk4 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20180.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 75; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
```

```
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk4 gene.
XX
SQ   Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;

Alignment Scores:
Pred. No.:                1.12e-42        Length:          580
Score:                    434.00          Matches:         75
Percent Similarity:       74.22%          Conservative:    20
Best Local Similarity:    58.59%          Mismatches:      33
Query Match:              60.70%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-2 (1-129) x AAD32057 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||  |||||||||   :::|||:::   :::||||||:::|||   ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              :::||||||   |||||||||   ||||||::::::|||       ||||||||||||||
Db         96 AGGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTC 155

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||   |||||||||   ||||||   ||||||:::|||:::   ||||||
Db        156 TCTATTGAAAGCGACGGAGAAGCAGCCTTTGTGGCCCAGCTGGTCGCTGAGAACATCAAG 215

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 |||      ||||||||||||||||||   |||   :::||||||||||||||||
Db        216 CAAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGGCGAAGAGAAGCAATGCAGC 275

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||||:::::   ||||||   |||:::      :::|||
Db        276 ACGAAGTGGAGCGATGGCTCCAGCGTCAATTATGAGAACCTGATTAAACATGCGACCAAA 335

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   ||:::      ||||||||||||      |||   |||   :::|||
Db        336 AAGTGTTTTGGGCTGAAAAAAGAGACAGGGTTTCGCACGTGGCGCAATGTTCACTGTACA 395

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||:::   |||:::|||:::
Db        396 CAACAAAATCTTTTCATGTGCAAG 419

RESULT 6
AAD32056
ID   AAD32056 standard; DNA; 456 BP.
XX
AC   AAD32056;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; ds.
XX
OS   Sistrurus miliarius.
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20179.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Disclosure; Page 74; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein Zsnk3, degenerate nucleic acid.
XX
SQ   Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

Alignment Scores:
Pred. No.:                2.23e-38        Length:          456
Score:                    397.00          Matches:         66
Percent Similarity:       59.69%          Conservative:    11
Best Local Similarity:    51.16%          Mismatches:      52
Query Match:              55.52%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-2 (1-129) x AAD32056 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||          |||      |||:::   |||||||||   ||||||||||||
Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   |||||||||   ||||||   :::|||       |||||||||   |||
Db        130 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNGARGGNGGNCAYYTNGTN 189

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||          ||||||   ||||||   |||   :::   :::
Db        190 WSNATHGARWSNWSNGARGARGCNGCNTTYGTNGCNCARYTNGTNCCNGARAAYMGNMGN 249

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||      ::::::||||||||         |||   ||||||||||||||||
Db        250 MGNGCNATHYTNTAYATHTGGATHGGNYTNMGNGTNCARGGNAARGARAARCARTGYWSN 309

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      :::      ||||||||||||||   |||   |||
Db        310 GCNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYTGGATHGARGCNGARWSNAAR 369

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||   |||              |||   :::|||||||||   |||   ||||||
Db        370 ACNTGYYTNGGNYTNCARCARGGNACNAAYTAYCAYAARTGGGTNAAYATHTAYTGYGGN 429

Qy        121 GlnGlnAspProPheValCysGluAla 129
              :::   :::||||||||||||||||||
```

```
Db        430 GARATHAAYCCNTTYGTNTGYGARGCN 456

RESULT 7
AAD32059
ID   AAD32059 standard; DNA; 725 BP.
XX
AC   AAD32059;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk5 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk5; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             88..564
FT                   /*tag=  a
FT                   /product= "Zsnk5 protein"
FT   sig_peptide     88..156
FT                   /*tag=  b
FT   mat_peptide     157..561
FT                   /*tag=  c
FT                   /product= "Mature Zsnk5 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20181.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 77; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk5 gene.
XX
SQ   Sequence 725 BP; 179 A; 171 C; 200 G; 175 T; 0 other;

Alignment Scores:

Pred. No.:                6.62e-37      Length:          725
Score:                    387.00        Matches:         67
Percent Similarity:       71.88%        Conservative:    25
Best Local Similarity:    52.34%        Mismatches:      36
Query Match:              54.13%        Indels:          0
DB:                       24            Gaps:            0

US-09-938-114-2 (1-129) x AAD32059 (1-725)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||   :::|||:::   :::||||||:::|||   |||:::
Db        163 AATTGTCCCTCTGGTTGGTTCGCCTACGATCAGTATTGCTACAGGGTCATCAAACGACTC 222

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   |||||||||   ||||||::::::|||      ||||||||||||
Db        223 AAGACCTGGGACGATGCAGAGCGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCG 282

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||:::       ||||||   |||:::   ||||||:::|||       ||||||
Db        283 TCTGTCGAAAACGATGAAGAAGCAGTCTTTCTGGCCCAGTTGGTCGCTGCGAACATAAAG 342

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                    :::   :::||||||||||||||||||   |||||||||   :::||||||||||
Db        343 CAAAACCAATACTATGTCTGGATTGGACTGAGGATTCAAAACAAAGGACAGCAATGCAGC 402

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||||:::|||   ||||||   ::::::      ||||||
Db        403 ACGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACCTGGTTAAATCACATTCCAAA 462

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||:::      ||||||   |||   :::|||   |||      ||||||
Db        463 AAGTGTTTTGGGCTGAAAAAAGAGACAGAGTTTCTTCAATGGTACAATACTGACTGCGAA 522

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::::::::   |||||||||:::
Db        523 GAAAAAAACCTTTTCGTCTGCAAG 546

RESULT 8
AAD32058
ID   AAD32058 standard; DNA; 432 BP.
XX
AC   AAD32058;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; ds.
XX
OS   Sistrurus miliarius.
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
```

PT molecular biology -
XX
PS Disclosure; Page 76; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk4, degenerate nucleic acid.
XX
SQ Sequence 432 BP; 91 A; 34 C; 84 G; 55 T; 168 other;

Alignment Scores:

Pred. No.:	3.96e-36	Length:	432
Score:	378.00	Matches:	64
Percent Similarity:	62.50%	Conservative:	16
Best Local Similarity:	50.00%	Mismatches:	48
Query Match:	52.87%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-2 (1-129) x AAD32058 (1-432)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||   |||||||||   ||||||   :::|||      |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGTN 153

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||      |||||||||   ||||||   |||   :::|||:::   ||||||
Db        154 WSNATHGARWSNGAYGGNGARGCNGCNTTYGTNGCNCARYTNGTNGCNGARAAYATHAAR 213

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||      ||||||||||||          |||   :::||||||||||||
Db        214 CARAAYAARTAYGAYGTNTGGATHGGNYTNMGNATHCARGGNGARGARAARCARTGYWSN 273

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      ::::::   ||||||   |||:::      :::|||
Db        274 ACNAARTGGWSNGAYGGNWSNWSNGTNAAYTAYGARAAYYTNATHAARCAYGCNACNAAR 333

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||         ||||||||||||      |||   |||   :::|||
Db        334 AARTGYTTYGGNYTNAARAARGARACNGGNTTYMGNACNTGGMGNAAYGTNCAYTGYACN 393

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||:::   |||:::|||:::
Db        394 CARCARAAYYTNTTYATGTGYAAR 417
```

RESULT 9
AAD32060
ID AAD32060 standard; DNA; 474 BP.
XX
AC AAD32060;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk5, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 78; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk5, degenerate nucleic acid.
XX
SQ Sequence 474 BP; 89 A; 31 C; 84 G; 74 T; 196 other;

Alignment Scores:

Pred. No.:	3.41e-30	Length:	474
Score:	329.00	Matches:	55
Percent Similarity:	58.59%	Conservative:	20
Best Local Similarity:	42.97%	Mismatches:	53
Query Match:	46.01%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-2 (1-129) x AAD32060 (1-474)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          |||   :::|||:::   :::||||||   |||   |||
Db         76 AAYTGYCCNWSNGGNTGGTTYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARMGNYTN 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   |||||||||   ||||||   :::|||      |||||||||
Db        136 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGCN 195

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              :::|||:::      ||||||   |||      |||   :::|||      ||||||
Db        196 WSNGTNGARAAYGAYGARGARGCNGTNTTYYTNGCNCARYTNGTNGCNGCNAAYATHAAR 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              :::   :::||||||||||||          |||||||||   ::|||||
Db        256 CARAAYCARTAYTAYGTNTGGATHGGNYTNMGNATHCARAAYAARGGNCARCARTGYWSN 315
```

```
Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||     :::     ||||||   ::::::          |||
Db        316 ACNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYYTNGTNAARWSNCAYWSNAAR 375

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||          ||||||   |||   :::|||   |||      ||||||
Db        376 AARTGYTTYGGNYTNAARAARGARACNGARTTYYTNCARTGGTAYAAYACNGAYTGYGAR 435

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::::::::   |||||||||:::
Db        436 GARAARAAYYTNTTYGTNTGYAAR 459

RESULT 10
AAI71877
ID   AAI71877 standard; DNA; 633 BP.
XX
AC   AAI71877;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin B chain coding sequence.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis; ds.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   P-PSDB; AAM51544.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 1; Page 11; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   encodes the B chain of halyxin.
XX
SQ   Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;

Alignment Scores:
Pred. No.:                2.41e-25        Length:         633
Score:                    290.00          Matches:        59
Percent Similarity:       57.81%          Conservative:   15
Best Local Similarity:    46.09%          Mismatches:     46
Query Match:              40.56%          Indels:         8
DB:                       22              Gaps:           4

US-09-938-114-2 (1-129) x AAI71877 (1-633)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||  ||||||||||||||||||||||||||||||   |||   :::
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAACCATTTAATGAACAA 152

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||   |||   |||||||||:::||||||||||:::|||      ||||||||||||||
Db        153 AAAAATTGGGCCGATGCAGAGAATTTCTGCACACAACAGCACACAGGCGGACATCTGGTC 212

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||       |||:::   ||||||||||||||   :::|||       |||
Db        213 TCCTTCCACAGCACTGAAGAAGCAGATTTTGTGGTCAAGCTGGCCTTCCAAAATTTTGGC 272

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                                  |||:::||||||      |||       |||||||||
Db        273 CACGGTATTTTC------TGGATGGGACTG------AGCAATGTCTGGAATCAATGCAGC 320

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||      :::   :::       |||   |||   ||||||
Db        321 TGGCAATGGAGCAGTGCTGCCAAGCTCAAATACGAAGCCTGGGCTGAAGAA------TCT 374

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||      |||:::   ::::::      :::|||||   :::      |||
Db        375 TATTGT------GTCTATTTCAAGTCAACTAATAACAAATGGAGGAGTAGAGCCTGCAGA 428

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::      ||||||||||||
Db        429 ATGGAGGCATATTTCGTCTGCGAG 452

RESULT 11
AAQ89309
ID   AAQ89309 standard; cDNA; 690 BP.
XX
AC   AAQ89309;
XX
DT   25-MAR-2003  (updated)
DT   28-NOV-1995  (first entry)
XX
DE   Snake venom antithrombotic oligopeptide cDNA.
XX
KW   Antithrombotic peptide; snake venom; platelet binding inhibition;
KW   von Willebrand factors; Crotalus horridus horridus; ds.
XX
OS   Crotalus horridus horridus.
XX
FH   Key             Location/Qualifiers
FT   CDS             66..515
FT                   /*tag=  a
XX
PN   WO9508573-A1.
XX
PD   30-MAR-1995.
XX
PF   21-SEP-1994;   94WO-JP01555.
XX
PR   22-SEP-1993;   93JP-0236975.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kito M,  Kobayashi T,  Nagano M;
PI   Tanaka A,  Yamamoto H,  Yoshimoto R;
XX
DR   WPI; 1995-139559/18.
DR   P-PSDB; AAR71981.
XX
PT   Single-chain antithrombotic peptide - obtained by cleaving an
PT   oligopeptide from snake venom to break inter-chain di:sulphide
PT   bonds but preserve intra-chain di:sulphide bonds
XX
PS   Example 2; Pages 47-48; 84pp; Japanese.
XX
CC   AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
CC   specifically from the snake venom oligopeptide AAR71981, encoded by
CC   AAQ89309. These peptides have the advantage of avoiding significant
CC   thrombocytopenia when administered at the minimun dose, for in vivo
CC   inhibition of platelet von Willebrand factor binding.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.:                7.13e-25        Length:         690
Score:                    286.50          Matches:        55
```

```
Percent Similarity:     60.16%          Conservative:  22
Best Local Similarity:  42.97%          Mismatches:    42
Query Match:            40.07%          Indels:        9
DB:                     16              Gaps:          2

US-09-938-114-2 (1-129) x AAQ89309 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   ||||||||||||:::   :::||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
                ||||||   |||||||||   ||||||::::::|||      ||||||||||||:::
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 260

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||::::::   ||||||   ||||||      ::::::            :::
Db        261 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 320

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                  ::::::||||||||||||||||   |||||||||   :::   |||
Db        321 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGC--- 377

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                           ||||||||||||   ||||||   :::      :::
Db        378 ------------------TCCAGCATCAGTTATGAGAACCTGGTT------GACCCATTT 413

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              :::|||      |||      :::|||         :::|||                |||||
Db        414 GAATGTTTTATGGTGAGCAGAGACACAAGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||      |||:::|||:::
Db        474 CAACAACATTCTTTCATATGCAAG 497

RESULT 12
AAC61144
ID   AAC61144 standard; DNA; 690 BP.
XX
AC   AAC61144;
XX
DT   07-FEB-2001  (first entry)
XX
DE   DNA encoding a snake venom derived protein.
XX
KW   Subunit peptide production; snake venom; rattlesnake; thrombolytic;
KW   von Willebrand's factor; blood platelet-inhibitory activity; ds.
XX
OS   Crotalus horridus horridus.
XX
PN   WO200059926-A1.
XX
PD   12-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-JP02127.
XX
PR   02-APR-1999;   99JP-0096073.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Fukuchi N,  Kageyama S,  Kito M,  Kayahara T,  Yamamoto H;
XX
DR   WPI; 2000-664985/64.
DR   P-PSDB; AAY85628.
XX
PT   Producing physiologically-active subunit peptides originating in
PT   polymer proteins by denaturation and specific separation, with lower
PT   antigenicity but improved solubility and stability, e.g. blood
PT   platelet-binding inhibitors -
XX
PS   Disclosure; Page 46; 51pp; Japanese.
XX
CC   This invention relates to a method for the production of a subunit
CC   peptide originating from a polymer protein with disulphide bonds within
CC   and between subunits. The method comprises denaturing the protein or its
CC   subunit using a protein denaturing agent in a solution, removing the
CC   agent in the presence of a polyoxyalkyl polyether which reacts with a
CC   thiol group and unwinds the subunit, and separating the polyoxyalkyl
CC   polyether-bound subunit peptide. The method can be used for producing
CC   physiologically-active subunit peptides for polymer proteins e.g. snake
CC   venom-originated dimer peptide with blood platelet-inhibitory activity
CC   on von Willebrand's factor. The peptides produced have platelet-binding
CC   inhibitory, and thrombolytic activity. The present sequence represents
CC   DNA encoding rattlesnake protein used in an example illustrating the
CC   method of the invention.
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.:              7.13e-25        Length:        690
Score:                  286.50          Matches:       55
Percent Similarity:     60.16%          Conservative:  22
Best Local Similarity:  42.97%          Mismatches:    42
Query Match:            40.07%          Indels:        9
DB:                     21              Gaps:          2

US-09-938-114-2 (1-129) x AAC61144 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   ||||||||||||:::   :::||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
                ||||||   |||||||||   ||||||::::::|||      ||||||||||||:::
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 260

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||::::::   ||||||   ||||||      ::::::            :::
Db        261 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 320

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                  ::::::||||||||||||||||   |||||||||   :::   |||
Db        321 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGC--- 377

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                           ||||||||||||   ||||||   :::      :::
Db        378 ------------------TCCAGCATCAGTTATGAGAACCTGGTT------GACCCATTT 413

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              :::|||      |||      :::|||         :::|||                |||||
Db        414 GAATGTTTTATGGTGAGCAGAGACACAAGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||      |||:::|||:::
Db        474 CAACAACATTCTTTCATATGCAAG 497

RESULT 13
AAK99834
ID   AAK99834 standard; DNA; 690 BP.
XX
AC   AAK99834;
XX
DT   19-JUL-2002  (first entry)
XX
DE   DNA encoding the antithrombotic wild-type rattlesnake protein.
XX
KW   Thrombolytic; glycoprotein Ib-binding protein; snake venom; rattlesnake;
KW   antithrombotic; glycoprotein Ib; long half life; low antigenicity;
KW   drug efficacy; gene; ds.
XX
OS   Crotalus horridus horridus.
XX
FH   Key             Location/Qualifiers
FT   CDS             66..515
FT                   /*tag=  a
FT                   /product= "Antithrombotic wild-type rattlesnake protein"
```

```
XX
PN   EP1195384-A1.
XX
PD   10-APR-2002.
XX
PF   04-OCT-2001; 2001EP-0123277.
XX
PR   04-OCT-2000; 2000JP-0305279.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Fukuchi N,  Kito M,  Kayahara T,  Futaki F,  Ishikawa K,  Suzuki E;
PI   Gondoh K,  Shimba N,  Yamada N;
XX
DR   WPI; 2002-364482/40.
DR   P-PSDB; AAO20974.
XX
PT   Glycoprotein Ib-binding protein, useful for treating thrombosis,
PT   comprises specific mutations in protein originating from snake venom  -
XX
PS   Example 1; Page 25; 49pp; English.
XX
CC   The invention relates to a glycoprotein Ib-binding protein, originating
CC   from snake venom, comprising specific mutations and antithrombotic
CC   activity. Glycoprotein Ib-binding protein is used in a drug having
CC   antithrombotic activity. Glycoprotein Ib-binding protein has a high
CC   binding activity to glycoprotein Ib, a long half life/drug efficacy
CC   retention in blood, and low antigenicity. This polynucleotide sequence
CC   represents DNA of the wild-type rattlesnake protein of the invention.
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.:                7.13e-25        Length:         690
Score:                    286.50          Matches:        55
Percent Similarity:       60.16%          Conservative:   22
Best Local Similarity:    42.97%          Mismatches:     42
Query Match:              40.07%          Indels:         9
DB:                       24              Gaps:           2

US-09-938-114-2 (1-129) x AAK99834 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||||||||||:::   :::|||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||   ||||||||   ||||||:::::|||       ||||||||||:::
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 260

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||::::::   ||||||   ||||||      ::::::          :::
Db        261 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 320

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                     ::::::|||||||||||||||   |||||||||   :::   |||
Db        321 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGC--- 377

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                           ||||||||||||   ||||||  :::     :::
Db        378 ------------------TCCAGCATCAGTTATGAGAACCTGGTT------GACCCATTT 413

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              :::|||       |||      :::|||        :::|||          ||||||
Db        414 GAATGTTTTATGGTGAGCAGAGACACAAGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||      |||:::|||:::
Db        474 CAACAACATTCTTTCATATGCAAG 497

RESULT 14
AAL42016
ID   AAL42016 standard; cDNA; 583 BP.
XX
AC   AAL42016;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Korean adder snake venom salmorin B chain protein cDNA sequence.
XX
KW   Korean adder; ss; salmorin protein B chain; snake venom;
KW   fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW   thrombin binding; blood coagulation.
XX
OS   Agkistrodon halys brevicaudus.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..444
FT                   /*tag=  a
FT                   /product= "Salmorin B chain protein"
FT   sig_peptide     7..75
FT                   /*tag=  b
FT                   /note= "Signal peptide"
FT   mat_peptide     76..441
FT                   /*tag=  c
FT                   /note= "Mature Salmorin B chain protein"
FT   3'UTR           445..570
FT                   /*tag=  d
FT   polyA_signal    571..583
FT                   /*tag=  e
XX
PN   WO200214514-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-KR01277.
XX
PR   27-JUL-2000; 2000KR-0043470.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Chung K,  Kim D,  Koh Y;
XX
DR   WPI; 2002-241907/29.
DR   P-PSDB; AAO14521.
XX
PT   New salmorin protein derived from venom of Korean adder Agkistrodon
PT   halys brevicaudus, useful for treating thrombosis by repressing
PT   fibrinogen clotting through repression of activation of prothrombin
PT   into thrombin  -
XX
PS   Claim 1; Fig 1B; 30pp; English.
XX
CC   The invention comprises the nucleotide and protein sequences of a
CC   salmorin protein derived from the venom of Korean adder. Salmorin protein
CC   is composed of an A chain and a B chain, and has inhibitory activity
CC   against fibrinogen clotting. The salmorin protein of the invention is
CC   useful for treating thrombosis, as it represses fibrinogen clotting
CC   potently by binding to prothrombin and thrombin so as to delay blood
CC   coagulation. The present cDNA sequence encodes the Korean adder salmorin
CC   B chain protein.
XX
SQ   Sequence 583 BP; 151 A; 137 C; 150 G; 145 T; 0 other;

Alignment Scores:
Pred. No.:                2.72e-23        Length:         583
Score:                    272.50          Matches:        56
Percent Similarity:       54.69%          Conservative:   14
Best Local Similarity:    43.75%          Mismatches:     49
Query Match:              38.11%          Indels:         9
DB:                       24              Gaps:           3

US-09-938-114-2 (1-129) x AAL42016 (1-583)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   |||||||||||||||||||||||||||||:::|||   |||
```

```
Db         76 GATTGTCCCTCTGGTTGGTCCTCTTATGAAGGGCATTGCTACAAGCTCTTCAATCAACAG 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        136 TCTCAGTGGGCCCATGCAGAGAAATTCTGCACACTACAGCACACAGGCGGACATCTGGTC 195

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        196 TCCTTCCACAGTACTGAAGAAGCAGATTTTGTGGTCAAGCTGGCCTTCCAAAATTTTGGC 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        256 CACGGTATTTTC------TGGATGGGACTG------AGCAATGTCTGGAATCAATGCAGC 303

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        304 TGGCAATGGAGCAATGCTGCCAAGCTCAAATACGAAGCCTGGGCTGAG------------ 351

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        352 ---ATCTATTGTGTCTATTTCAAGTCAACTAATAACAAATGGAGGAGTAGAGCCTGCAGA 408

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        409 ATGGAGGCATATTTCGTCTGCGAG 432
```

```
RESULT 15
AAD32053
ID   AAD32053 standard; DNA; 483 BP.
XX
AC   AAD32053;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk2 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk2; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..458
FT                   /*tag=  a
FT                   /product= "Zsnk2 protein"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   sig_peptide     3..59
FT                   /*tag=  b
FT   mat_peptide     60..455
FT                   /*tag=  c
FT                   /product= "Mature Zsnk2 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20178.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 71-72; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk2 gene.
XX
SQ   Sequence 483 BP; 129 A; 112 C; 123 G; 119 T; 0 other;
```

Alignment Scores:			
Pred. No.:	2.21e-22	Length:	483
Score:	264.00	Matches:	55
Percent Similarity:	57.03%	Conservative:	18
Best Local Similarity:	42.97%	Mismatches:	49
Query Match:	36.92%	Indels:	6
DB:	24	Gaps:	3

US-09-938-114-2 (1-129) x AAD32053 (1-483)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        120 AAAACCTGGGATGATGCAGAGAGTTTCTGCTACACACAGCACAGAGACAGCCGCCTGGCC 179

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        180 TCCATCCACAGCAGTGAAGAAGAAGCTTTTGTGGGCAAACTGGCCTCCCAAACTTTGAAA 239

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        240 TTCACTTCC------ATGTGGATCGGACTGAAA------GATCTATGGAAAGAATGCAAA 287

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        288 TGGCAGTGGAGCGATGACACCAAACTGGACTACAAAGCCTGGACTCGAAGA------CCC 341

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        342 TATTGTACAGTAATGGTAGTCAAGACAGATAGGATCTTTTGGTTCAATAGAGGTTGCGAA 401

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        402 AAGACTGTATCTTTTGTCTGCAAG 425
```

```
Search completed: December  8, 2003, 14:04:46
Job time : 446.023 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:49:04 ; Search time 101.726 Seconds
(without alignments)
559.725 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091001_22878/app_query.fasta_1.725
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_95_@runat_08122003_091001_22878 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	653	91.3	544	4	US-09-058-740-1	Sequence 1, Appli
2	286.5	40.1	690	2	US-08-612-840A-7	Sequence 7, Appli
3	216	30.2	454	4	US-09-058-740-12	Sequence 12, Appl
4	201.5	28.2	777	3	US-09-146-969-1	Sequence 1, Appli
5	201.5	28.2	4771	3	US-08-840-062-3	Sequence 3, Appli
6	197.5	27.6	4588	3	US-08-840-062-1	Sequence 1, Appli
7	195.5	27.3	590	2	US-08-454-557C-32	Sequence 32, Appl
8	195.5	27.3	590	2	US-08-340-426D-32	Sequence 32, Appl
9	195.5	27.3	590	2	US-08-450-673C-32	Sequence 32, Appl
10	195.5	27.3	590	5	PCT-US95-17111A-32	Sequence 32, Appl
11	184.5	25.8	5191	1	US-08-340-428B-1	Sequence 1, Appli
12	184.5	25.8	5191	5	PCT-US93-07306-1	Sequence 1, Appli
13	183	25.6	522	1	US-07-778-156-4	Sequence 4, Appli
14	183	25.6	522	2	US-08-422-166-4	Sequence 4, Appli
15	183	25.6	797	2	US-08-464-637-1	Sequence 1, Appli
16	183	25.6	797	2	US-08-822-261-5	Sequence 5, Appli
17	183	25.6	797	4	US-09-226-852-5	Sequence 5, Appli
18	183	25.6	798	1	US-07-778-156-12	Sequence 12, Appl
19	183	25.6	798	2	US-08-822-261-6	Sequence 6, Appli
20	183	25.6	798	2	US-08-422-166-12	Sequence 12, Appl
21	183	25.6	798	3	US-09-146-969-2	Sequence 2, Appli
22	183	25.6	798	4	US-09-226-852-6	Sequence 6, Appli
23	181.5	25.4	272	2	US-08-612-840A-6	Sequence 6, Appli
24	181.5	25.4	614	2	US-08-729-103-2	Sequence 2, Appli
25	181.5	25.4	1114	2	US-08-468-413-1	Sequence 1, Appli
26	181.5	25.4	1114	3	US-09-162-508-1	Sequence 1, Appli
27	181.5	25.4	1114	5	PCT-US95-07169-1	Sequence 1, Appli
28	175	24.5	558	1	US-08-909-725-4	Sequence 4, Appli
29	175	24.5	586	3	US-09-146-969-3	Sequence 3, Appli
30	175	24.5	747	2	US-08-401-530A-1	Sequence 1, Appli
31	175	24.5	747	2	US-08-709-662-1	Sequence 1, Appli
32	175	24.5	762	2	US-08-822-261-2	Sequence 2, Appli
33	175	24.5	762	4	US-09-226-852-2	Sequence 2, Appli
34	168	23.5	4978	1	US-08-220-603A-1	Sequence 1, Appli
35	163	22.8	885	1	US-08-365-103B-3	Sequence 3, Appli
36	163	22.8	924	1	US-08-365-103B-5	Sequence 5, Appli
37	163	22.8	1005	1	US-08-365-103B-1	Sequence 1, Appli
38	161.5	22.6	8224	6	5180808-1	Patent No. 5180808
39	159	22.2	474	1	US-07-778-156-8	Sequence 8, Appli
40	159	22.2	474	2	US-08-422-166-8	Sequence 8, Appli
41	159	22.2	793	1	US-07-778-156-1	Sequence 1, Appli
42	159	22.2	793	2	US-08-422-166-1	Sequence 1, Appli
43	157.5	22.0	3259	5	PCT-US95-03747-1	Sequence 1, Appli
44	155	21.7	5169	4	US-09-194-612A-2	Sequence 2, Appli
45	144	20.1	2318	4	US-09-620-312D-733	Sequence 733, App

ALIGNMENTS

```
RESULT 1
US-09-058-740-1
; Sequence 1, Application US/09058740
; Patent No. 6489451
;  GENERAL INFORMATION:
;       APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                  Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                  Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                  Jin-Guo Ding, Fang Rong, Yan Liu and
;                  Hui-Ran Chen
;       TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                           VENOM OF AGKISTRODON ACUTUS
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/058,740
;            FILING DATE: 10-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: <Unknown>
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Chen, Anthony C.
```

```
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 544 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-058-740-1

Alignment Scores:
Pred. No.:              8.38e-75        Length:         544
Score:                  653.00          Matches:        119
Percent Similarity:     99.17%          Conservative:   0
Best Local Similarity:  99.17%          Mismatches:     1
Query Match:            91.33%          Indels:         0
DB:                     4               Gaps:           0

US-09-938-114-2 (1-129) x US-09-058-740-1 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy         30 CysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAsp 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy         50 PheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGly 69
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy         70 LeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIle 89
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy         90 SerLysGluAsnTrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThr 109
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy        110 GlyPheHisLysTrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGluAla 129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

RESULT 2
US-08-612-840A-7
; Sequence 7, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  690 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Crotalus horridus horridus
;      STRAIN:
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  66..512
US-08-612-840A-7

Alignment Scores:
Pred. No.:              1.14e-27        Length:         690
Score:                  286.50          Matches:        55
Percent Similarity:     60.16%          Conservative:   22
Best Local Similarity:  42.97%          Mismatches:     42
Query Match:            40.07%          Indels:         9
DB:                     2               Gaps:           2

US-09-938-114-2 (1-129) x US-08-612-840A-7 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||  ||||||||||||:::  :::|||||||||  |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||  |||||||||  |||||:::::|||      ||||||||||:::
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 260

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||::::::   ||||||  ||||||      ::::::           :::
Db        261 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 320

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                    ::::::||||||||||||||  ||||||||  :::   |||
Db        321 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGC--- 377

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                           ||||||||||||  ||||||  :::     :::
Db        378 -----------------TCCAGCATCAGTTATGAGAACCTGGTT------GACCCATTT 413

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              :::|||      |||      :::|||          :::|||              ||||||
Db        414 GAATGTTTTATGGTGAGCAGAGACACAAGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||      |||:::|||:::
Db        474 CAACAACATTCTTTCATATGCAAG 497

RESULT 3
```

```
US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
;   GENERAL INFORMATION:
;        APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                   Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                   Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                   Jin-Guo Ding, Fang Rong, Yan Liu and
;                   Hui-Ran Chen
;        TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                            VENOM OF AGKISTRODON ACUTUS
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/058,740
;             FILING DATE: 10-Apr-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Chen, Anthony C.
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 454 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        FEATURE:
;             NAME/KEY:          Coding Sequence
;             LOCATION:          3...440
;             OTHER INFORMATION:   "N" stands for any base.
;             "Xaa" stands for any amino acid.
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-058-740-12

Alignment Scores:
Pred. No.:                6.87e-19        Length:          454
Score:                    216.00          Matches:         50
Percent Similarity:       49.63%          Conservative:    17
Best Local Similarity:    37.04%          Mismatches:      46
Query Match:              30.21%          Indels:          22
DB:                       4               Gaps:            6

US-09-938-114-2 (1-129) x US-09-058-740-12 (1-454)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||:::|||||||||||||||||||||||||||||||   |||   :::
Db         72 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 131

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeu--- 39
              ||||||||   ||||||||   |||||||||:::|||      |||   ||||||
Db        132 AAGACCTGGGCAGATGCAGAGAAATTCTGCACACAACAACACAAAGGCAGCCATCTGCCT 191

Qy         40 ------------------ValSerIleGluSerSerGlyGluAlaAspPheValGlyGln 53
                              :::                 |||      |||
Db        192 CTCACAGCAGTGAGAGCGATTGTGTNNNNNNNNNNTGGTCACGTTGACCACACC------ 245

Qy         54 LeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGln 73
                         |||:::|||          :::          ||||||||::::
Db        246 ------------AAGTTGAAACTGATTAGTCTG---------ATTGGACTGAAG------ 278

Qy         74 AsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn 93
              |||           |||     :::||||||||||||:::   :::      ::::::
Db        279 AACATCTGGAACGGATGCTACTGGAAGTGGAGCGATGGCACCAAGCTCGACTACAAAGAC 338

Qy         94 TrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
              |||         |||      :::||||||                |||      ::::::
Db        339 TGG------CGTGAACAATTTGAATGTCTC------GTATCCAGGACAGTTAATAACGAA 386

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||   :::     |||               |||||||||:::
Db        387 TGGCTAAGTATGGACTGCGGCACTACTTGCTCTTTCGTCTGCAAG 431

RESULT 4
US-09-146-969-1
; Sequence 1, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
;  APPLICANT: Dieckgraefe, Brian K.
;  TITLE OF INVENTION: Gene Markers for Chronic Mucosal Injury
;  FILE REFERENCE: 04255.75314
;  CURRENT APPLICATION NUMBER: US/09/146,969
;  CURRENT FILING DATE: 1998-09-04
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 777
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-146-969-1

Alignment Scores:
Pred. No.:                1.07e-16        Length:          777
Score:                    201.50          Matches:         44
Percent Similarity:       53.03%          Conservative:    26
Best Local Similarity:    33.33%          Mismatches:      53
Query Match:              28.18%          Indels:          9
DB:                       3               Gaps:            5

US-09-938-114-2 (1-129) x US-09-146-969-1 (1-777)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||          ::::::|||      :::||||||         :::      :::
Db        188 TGCCCAGAAGGCACCAATGCCTATCGCTCCTACTGCTACTACTTTAATGAAGACCGTGAG 247

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              ||||||   ||||||::   :::|||   :::   :::|||   |||:::||||||||||
Db        248 ACCTGGGTTGATGCAGATCTCTATTGC---CAGAACATGAATTCGGGCAACCTGGTGTCT 304

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   :::     |||     ||||||     ||||||  :::      :::
Db        305 GTGCTCACCCAGGCCGAGGGTGCCTTTGTGGCCTCACTGATTAAGGAGAGT---GGCACT 361

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
                    :::|||||||||||||||           |||   ::::::
Db        362 GATGACTTCAATGTCTGGATTGGCCTCCATGACCCCAAAAAGAACCGCCGCTGG------ 415

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrp---IleGluGluGluSerLys 100
              ||||||   ||||||   :::|||   ::::::|||   |||          |||
Db        416 CACTGGAGCAGTGGGTCCCTGGTCTCCTACAAGTCCTGGGGCATTGGAGCCCCAAGCAGT 475

Qy        101 ------------LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsn 116
                          |||:::   :::          |||||||||   ||||||::::::
Db        476 GTTAATCCTGGCTACTGTGTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGAT 535

Qy        117 PheTyrCysGluGlnGlnAspProPheValCysGlu 128
```

```
                    ||||||   :::      |||||||||:::
Db        536 GTGCCTTGTGAAGACAAGTTCTCCTTTGTATGCAAG 571

RESULT 5
US-08-840-062-3
; Sequence 3, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4771 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-840-062-3

Alignment Scores:
Pred. No.:                1.4e-15        Length:          4771
Score:                    201.50         Matches:         39
Percent Similarity:       52.59%         Conservative:    32
Best Local Similarity:    28.89%         Mismatches:      47
Query Match:              28.18%         Indels:          17
DB:                       3              Gaps:            5

US-09-938-114-2 (1-129) x US-08-840-062-3 (1-4771)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          |||      ::::::||||||||||||::::::      :::
Db       1153 GAGTGCGAGCCGAGCTGGCAGCCCTTCCAGGGCCACTGCTACCGCCTGCAGGCCGAGAAG 1212

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ::::::|||   :::::::::      |||   :::          ||||||   ||||||
Db       1213 CGCAGCTGGCAGGAGTCCAAGAAGGCATGTCTACGG------GGCGGTGGCGACCTGGTC 1266

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||   |||      |||   :::|||:::   :::   |||   |||:::::::::
Db       1267 AGCATCCACAGCATGGCGGAGCTGGAATTCATCACCAAGCAGATCAAGCAAGAGGTGGAG 1326

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                        :::||||||||||||||         :::   :::   |||   :::
Db       1327 ------------GAGCTGTGGATCGGCCTC------AACGATTTGAAGCTGCAGATGAAT 1368

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrp----------------- 94
                ||||||||||||||||||   :::|||      :::|||
Db       1369 TTTGAGTGGTCTGACGGGAGCCTTGTGAGCTTCACCCACTGGCACCCCTTTGAGCCCAAC 1428

Qy         95 ---IleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
                      :::      :::   |||:::   :::          |||      :::
Db       1429 AACTTCCGGGACAGTCTGGAGGACTGTGTCACCATCTGGGGCCCGGAAGGC------CGC 1482

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||   :::      |||   |||      |||   :::|||:::
Db       1483 TGGAACGACAGTCCCTGTAACCAGTCCTTGCCATCCATCTGCAAG 1527

RESULT 6
US-08-840-062-1
; Sequence 1, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4588 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-840-062-1

Alignment Scores:
Pred. No.:                4.32e-15       Length:          4588
Score:                    197.50         Matches:         38
Percent Similarity:       52.59%         Conservative:    33
Best Local Similarity:    28.15%         Mismatches:      47
Query Match:              27.62%         Indels:          17
DB:                       3              Gaps:            5

US-09-938-114-2 (1-129) x US-08-840-062-1 (1-4588)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          |||      ::::::||||||||||||::::::      :::
Db       1242 GAATGTGACCCCAGCTGGCAGCCCTTCCAGGGCCACTGCTACCGCCTGCAGGCCGAGAAG 1301

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ::::::|||   :::::::::      |||   :::          ||||||   |||:::
Db       1302 CGCAGCTGGCAGGAGTCCAAGAGGGCGTGTCTGCGG------GGTGGGGGTGACCTCCTT 1355

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||   |||      |||   :::|||:::   :::   |||   |||:::::::::
Db       1356 AGCATCCACAGCATGGCTGAGCTGGAGTTCATCACCAAACAGATCAAGCAAGAGGTGGAG 1415

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
```

```
                          :::|||||||||||||          :::   :::   |||   :::
Db       1416 ------------GAGCTATGGATTGGCCTC------AATGATTTGAAACTGCAGATGAAT 1457

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrp------------------ 94
                |||||||||||||||||   :::|||      :::|||
Db       1458 TTTGAGTGGTCCGACGGGAGCCTCGTGAGCTTCACCCACTGGCACCCCTTTGAGCCCAAC 1517

Qy         95 ---IleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
                    :::     :::  |||:::   :::           |||      :::
Db       1518 AACTTTCGTGACAGCCTGGAGGACTGTGTCACCATCTGGGGGCCGGAAGGA------CGC 1571

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||   :::      |||   |||      |||   :::|||:::
Db       1572 TGGAACGACAGTCCCTGTAACCAGTCCTTGCCATCCATTTGCAAG 1616

RESULT 7
US-08-454-557C-32
; Sequence 32, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-32

Alignment Scores:
Pred. No.:                4.24e-16        Length:          590
Score:                    195.50          Matches:         43
Percent Similarity:       53.97%          Conservative:    25
Best Local Similarity:    34.13%          Mismatches:      49
Query Match:              27.34%          Indels:          9
DB:                       2               Gaps:            5

US-09-938-114-2 (1-129) x US-08-454-557C-32 (1-590)

Qy          8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
              :::|||      :::||||||         :::      :::|||||   ||||||:::
Db          3 GCCTATCGCTCCTACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGAT 62

Qy         28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGlu 47
                 :::|||   :::   :::|||   |||:::||||||||||:::   :::       |||
Db         63 CTCTATTGC---CAGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAG 119

Qy         48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67
                 ||||||      ||||||   :::          :::           :::||||||
Db        120 GGTGCCTTTGTGGCCTCACTGATTAAGGAGAGT---GGCACTGATGACTTCAATGTCTGG 176

Qy         68 IleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySer 87
              |||||||||          |||   ::::::           ||||||   ||||||
Db        177 ATTGGCCTCCATGACCCCAAAAAGAACCGCCGCTGG------CACTGGAGCAGTGGGTCC 230

Qy         88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----------LysCys 102
                 :::|||   ::::::|||   |||          |||                  |||
Db        231 CTGGTCTCCTACAAGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGT 290

Qy        103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGlnGln 122
              :::   :::          |||||||||   ||||||::::::      ||||||   :::
Db        291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350

Qy        123 AspProPheValCysGlu 128
                |||||||||:::
Db        351 TTCTCCTTTGTCTGCAAG 368

RESULT 8
US-08-340-426D-32
; Sequence 32, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-32

Alignment Scores:
Pred. No.:                4.24e-16        Length:          590
Score:                    195.50          Matches:         43
Percent Similarity:       53.97%          Conservative:    25
Best Local Similarity:    34.13%          Mismatches:      49
Query Match:              27.34%          Indels:          9
DB:                       2               Gaps:            5

US-09-938-114-2 (1-129) x US-08-340-426D-32 (1-590)

Qy          8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
```

```
Db          3 GCCTATCGCTCCTACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGAT 62

Qy         28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGlu 47

Db         63 CTCTATTGC---CAGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAG 119

Qy         48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67

Db        120 GGTGCCTTTGTGGCCTCACTGATTAAGGAGAGT---GGCACTGATGACTTCAATGTCTGG 176

Qy         68 IleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySer 87

Db        177 ATTGGCCTCCATGACCCCAAAAAGAACCGCCGCTGG------CACTGGAGCAGTGGGTCC 230

Qy         88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----------LysCys 102

Db        231 CTGGTCTCCTACAAGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGT 290

Qy        103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGlnGln 122

Db        291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350

Qy        123 AspProPheValCysGlu 128

Db        351 TTCTCCTTTGTCTGCAAG 368

RESULT 9
US-08-450-673C-32
; Sequence 32, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-32

Alignment Scores:
Pred. No.:                4.24e-16        Length:          590
Score:                    195.50          Matches:         43
Percent Similarity:       53.97%          Conservative:    25
Best Local Similarity:    34.13%          Mismatches:      49
Query Match:              27.34%          Indels:          9
DB:                       2               Gaps:            5

US-09-938-114-2 (1-129) x US-08-450-673C-32 (1-590)

Qy          8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27

Db          3 GCCTATCGCTCCTACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGAT 62

Qy         28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGlu 47

Db         63 CTCTATTGC---CAGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAG 119

Qy         48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67

Db        120 GGTGCCTTTGTGGCCTCACTGATTAAGGAGAGT---GGCACTGATGACTTCAATGTCTGG 176

Qy         68 IleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySer 87

Db        177 ATTGGCCTCCATGACCCCAAAAAGAACCGCCGCTGG------CACTGGAGCAGTGGGTCC 230

Qy         88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----------LysCys 102

Db        231 CTGGTCTCCTACAAGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGT 290

Qy        103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGlnGln 122

Db        291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350

Qy        123 AspProPheValCysGlu 128

Db        351 TTCTCCTTTGTCTGCAAG 368

RESULT 10
PCT-US95-17111A-32
; Sequence 32, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  590 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
```

```
;      TOPOLOGY:  both
PCT-US95-17111A-32

Alignment Scores:
Pred. No.:                4.24e-16        Length:          590
Score:                    195.50          Matches:         43
Percent Similarity:       53.97%          Conservative:    25
Best Local Similarity:    34.13%          Mismatches:      49
Query Match:              27.34%          Indels:          9
DB:                       5               Gaps:            5

US-09-938-114-2 (1-129) x PCT-US95-17111A-32 (1-590)

Qy          8 SerTyrGluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGlu 27
              :::|||      :::||||||         :::      :::||||||   ||||||::: 
Db          3 GCCTATCGCTCCTACTGCTACTACTTTAATGAAGACCGTGAGACCTGGGTTGATGCAGAT 62

Qy         28 SerPheCysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGlu 47
                ::|||   :::   :::|||   |||:::||||||||:::   :::       |||
Db         63 CTCTATTGC---CAGAACATGAATTCGGGCAACCTGGTGTCTGTGCTCACCCAGGCCGAG 119

Qy         48 AlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrp 67
                ||||||      ||||||   :::       :::       :::||||||
Db        120 GGTGCCTTTGTGGCCTCACTGATTAAGGAGAGT---GGCACTGATGACTTCAATGTCTGG 176

Qy         68 IleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySer 87
              |||||||||          |||   ::::::          ||||||   ||||||
Db        177 ATTGGCCTCCATGACCCCAAAAAGAACCGCCGCTGG------CACTGGAGCAGTGGGTCC 230

Qy         88 SerIleSerLysGluAsnTrp---IleGluGluGluSerLys-----------LysCys 102
                :::|||   ::::::|||   |||          |||                |||
Db        231 CTGGTCTCCTACAAGTCCTGGGGCATTGGAGCCCCAAGCAGTGTTAATCCTGGCTACTGT 290

Qy        103 LeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGlnGln 122
                :::   :::         |||||||||   |||||::::::      ||||||   :::
Db        291 GTGAGCCTGACCTCAAGCACAGGATTCCAGAAATGGAAGGATGTGCCTTGTGAAGACAAG 350

Qy        123 AspProPheValCysGlu 128
                    |||||||||:::
Db        351 TTCTCCTTTGTCTGCAAG 368

RESULT 11
US-08-340-428B-1
; Sequence 1, Application US/08340428B
; Patent No. 5648465
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,428B
;      FILING DATE:  14 No. 5648465ember 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/922,911
;      FILING DATE:  03 August 1992
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  Margolis=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5191 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  77..3847
US-08-340-428B-1

Alignment Scores:
Pred. No.:                2.39e-13        Length:          5191
Score:                    184.50          Matches:         39
Percent Similarity:       45.93%          Conservative:    23
Best Local Similarity:    28.89%          Mismatches:      52
Query Match:              25.80%          Indels:          21
DB:                       1               Gaps:            6

US-09-938-114-2 (1-129) x US-08-340-428B-1 (1-5191)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||         |||       ::::::||||||||||||:::   |||         :::
Db       3161 TGCGACCGTGGCTGGCACAAATTCCAGGGCCACTGCTACCGCTACTTTGCTCATCGGCGG 3220

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
                |||   |||||||||      |||   ::::::         |||||||||   |||
Db       3221 GCCTGGGAGGACGCAGAGAGAGACTGCAGGCGCCGA------GCCGGCCACCTGACAAGT 3274

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   |||      |||      |||:::
Db       3275 GTCCACTCCCCAGAAGAGCACAAGTTTATTAACAGTTTT---------------------- 3313

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
                    :::   ||||||||||||       :::      |||:::
Db       3314 GGACACGAGAATTCATGGATTGGCCTGAATGACAGGACAGTAGAGAGG------GACTTC 3367

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGlu--------- 98
              :::|||:::|||   :::   :::      |||||||||   |||::::::
Db       3368 CAGTGGACAGACAACACAGGACTGCAATATGAGAACTGGAGAGAGAAGCAGCCGGATAAT 3427

Qy         99 -----------SerLysLysCys---LeuGlyValHisIleGluThrGlyPheHisLys 113
                         :::   |||   :::   |||      |||   |||      :::
Db       3428 TTCTTCGCAGGTGGGGAGGATTGTGTGGTGATGGTGGCGCATGAGAATGGA------CGC 3481

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||   :::      |||            |||:::||||||:::
Db       3482 TGGAATGATGTCCCCTGTAACTACAACCTCCCCTACGTCTGCAAG 3526

RESULT 12
PCT-US93-07306-1
; Sequence 1, Application PC/TUS9307306
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07306
;      FILING DATE:  03-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/922,911
;      FILING DATE:  03-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Guy K.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  Margolis=1A PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5191 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  77..3847
PCT-US93-07306-1

Alignment Scores:
Pred. No.:                2.39e-13        Length:         5191
Score:                    184.50          Matches:        39
Percent Similarity:       45.93%          Conservative:   23
Best Local Similarity:    28.89%          Mismatches:     52
Query Match:              25.80%          Indels:         21
DB:                       5               Gaps:           6

US-09-938-114-2 (1-129) x PCT-US93-07306-1 (1-5191)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||          |||      ::::::||||||||||||:::   |||         :::
Db       3161 TGCGACCGTGGCTGGCACAAATTCCAGGGCCACTGCTACCGCTACTTTGCTCATCGGCGG 3220

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
                ||| |||||||||      |||   ::::::        |||||||||   |||
Db       3221 GCCTGGGAGGACGCAGAGAGAGACTGCAGGCGCCGA------GCCGGCCACCTGACAAGT 3274

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   |||       |||       |||:::
Db       3275 GTCCACTCCCCAGAAGAGCACAAGTTTATTAACAGTTTT---------------------- 3313

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
                    :::   ||||||||||||       :::       |||:::
Db       3314 GGACACGAGAATTCATGGATTGGCCTGAATGACAGGACAGTAGAGAGG------GACTTC 3367

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGlu--------- 98
              :::|||:::|||   :::   :::       |||||||||   |||::::::
Db       3368 CAGTGGACAGACAACACAGGACTGCAATATGAGAACTGGAGAGAGAAGCAGCCGGATAAT 3427

Qy         99 ------------SerLysLysCys---LeuGlyValHisIleGluThrGlyPheHisLys 113
                        :::   |||   :::   |||       |||   |||       :::
Db       3428 TTCTTCGCAGGTGGGGAGGATTGTGTGGTGATGGTGGCGCATGAGAATGGA------CGC 3481

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||   :::       |||             |||:::||||||:::
Db       3482 TGGAATGATGTCCCCTGTAACTACAACCTCCCCTACGTCTGCAAG 3526

RESULT 13
US-07-778-156-4
; Sequence 4, Application US/07778156
; Patent No. 5436169
;  GENERAL INFORMATION:
;    APPLICANT:  IOVANNA, JUAN-LUCIO
;    APPLICANT:  KEIM, VOLKER
;    APPLICANT:  DAGORN, JEAN-CHARLES
;    TITLE OF INVENTION:  PROTEIN ASSOCIATED WITH ACUTE
;    TITLE OF INVENTION:  PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
;    TITLE OF INVENTION:  PANCREATITIS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/778,156
;      FILING DATE:  19911219
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5436169man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  354-012-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)521-4500
;      TELEFAX:  (703)486-2347
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  522 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  pancreas
US-07-778-156-4

Alignment Scores:
Pred. No.:                1.43e-14        Length:         522
Score:                    183.00          Matches:        43
Percent Similarity:       48.89%          Conservative:   23
Best Local Similarity:    31.85%          Mismatches:     59
Query Match:              25.59%          Indels:         10
DB:                       1               Gaps:           5

US-09-938-114-2 (1-129) x US-07-778-156-4 (1-522)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||                 :::|||      |||||||||   :::|||         |||
Db        118 TGTCCCAAAGGCTCCAAGGCCTATGGCTCCCACTGCTATGCCTTGTTTTTGTCACCAAAA 177

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              :::||||||||||:::      |||   |||:::   :::   |||:::|||||||||
Db        178 TCCTGGACAGATGCAGATCTGGCCTGCCAGAAGCGGCCCTCT---GGAAACCTGGTGTCT 234

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   |||       |||      ||||||      |||:::   :::   |||   :::
Db        235 GTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTG---AAGAGCATTGGTAAC 291

Qy         62 AlaLysIleHisValTrpIleGlyLeuArg------AlaGlnAsnLysGluLysGlnCys 79
              :::       :::|||||||||||||||          |||       |||
Db        292 AGCTACTCATACGTCTGGATTGGGCTCCATGACCCCACACAGGGCACCGAGCCCAATGGA 351

Qy         80 ---SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGlu 98
                 |||||||||             ::::::          |||
Db        352 GAAGGTTGGGAGTGGAGTAGCAGTGATGTGATGAATTACTTTGCATGGGAGAGAAATCCC 411
```

```
Qy         99 SerLys---------------LysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
              |||                  |||       :::         |||   |||   :::
Db        412 TCCACCATCTCAAGCCCCGGCCACTGTGCGAGCCTGTCGAGAAGCACAGCATTTCTGAGG 471

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||:::::::::   |||      :::   |||:::||||||:::
Db        472 TGGAAAGATTATAACTGTAATGTGAGGTTACCCTATGTCTGCAAA 516

RESULT 14
US-08-422-166-4
; Sequence 4, Application US/08422166
; Patent No. 5959086
;  GENERAL INFORMATION:
;    APPLICANT:  IOVANNA, JUAN-LUCIO
;    APPLICANT:  KEIM, VOLKER
;    APPLICANT:  DAGORN, JEAN-CHARLES
;    TITLE OF INVENTION:  PROTEIN ASSOCIATED WITH ACUTE
;    TITLE OF INVENTION:  PANCREATITIS, AGENTS FOR THE DIAGNOSIS OF ACUTE
;    TITLE OF INVENTION:  PANCREATITIS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/422,166
;      FILING DATE:  14-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/778,156
;      FILING DATE:  19-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5959086man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  354-012-0 PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)521-4500
;      TELEFAX:  (703)486-2347
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  522 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  pancreas
US-08-422-166-4

Alignment Scores:
Pred. No.:                1.43e-14        Length:          522
Score:                    183.00          Matches:         43
Percent Similarity:       48.89%          Conservative:    23
Best Local Similarity:    31.85%          Mismatches:      59
Query Match:              25.59%          Indels:          10
DB:                       2               Gaps:            5

US-09-938-114-2 (1-129) x US-08-422-166-4 (1-522)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||               :::|||     |||||||||  :::|||         |||

Db        118 TGTCCCAAAGGCTCCAAGGCCTATGGCTCCCACTGCTATGCCTTGTTTTTGTCACCAAAA 177

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              :::|||||||||||||:::     |||   |||:::   :::   |||:::|||||||||
Db        178 TCCTGGACAGATGCAGATCTGGCCTGCCAGAAGCGGCCCTCT---GGAAACCTGGTGTCT 234

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   |||       |||       ||||||       |||:::   :::   |||   :::
Db        235 GTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTG---AAGAGCATTGGTAAC 291

Qy         62 AlaLysIleHisValTrpIleGlyLeuArg------AlaGlnAsnLysGluLysGlnCys 79
              :::       :::|||||||||||||||||              |||       |||
Db        292 AGCTACTCATACGTCTGGATTGGGCTCCATGACCCCACACAGGGCACCGAGCCCAATGGA 351

Qy         80 ---SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGlu 98
                 |||||||||           ::::::          |||
Db        352 GAAGGTTGGGAGTGGAGTAGCAGTGATGTGATGAATTACTTTGCATGGGAGAGAAATCCC 411

Qy         99 SerLys---------------LysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
              |||                  |||       :::         |||   |||   :::
Db        412 TCCACCATCTCAAGCCCCGGCCACTGTGCGAGCCTGTCGAGAAGCACAGCATTTCTGAGG 471

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||:::::::::   |||      :::   |||:::||||||:::
Db        472 TGGAAAGATTATAACTGTAATGTGAGGTTACCCTATGTCTGCAAA 516

RESULT 15
US-08-464-637-1
; Sequence 1, Application US/08464637
; Patent No. 5834214
;  GENERAL INFORMATION:
;    APPLICANT:  Iovanna, Juan-Lucio
;    APPLICANT:  Dagorn, Jean-Charles
;    APPLICANT:  Keim, Volker
;    APPLICANT:  Sarles, Jacques
;    TITLE OF INVENTION:  Detection of Pancreatitis-Associated
;    TITLE OF INVENTION:  Protein for Diagnosis of Cystic Fibrosis or Pancreatic
;    TITLE OF INVENTION:  Disease (as amended).
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolasch & Birch, LLP
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-3487
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,637
;      FILING DATE:  30-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy Jr., Gerald M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  2121-107P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
```

```
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  43..567
;      OTHER INFORMATION:  /product= "human
;      OTHER INFORMATION:  pancreatitis-associated protein"
;      OTHER INFORMATION:  /note= "see, Fig. 3"
US-08-464-637-1

Alignment Scores:
Pred. No.:                2.6e-14        Length:          797
Score:                    183.00         Matches:         43
Percent Similarity:       48.89%         Conservative:    23
Best Local Similarity:    31.85%         Mismatches:      59
Query Match:              25.59%         Indels:          10
DB:                       2              Gaps:            5

US-09-938-114-2 (1-129) x US-08-464-637-1 (1-797)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||                :::|||      ||||||||  :::|||          |||
Db        160 TGTCCCAAAGGCTCCAAGGCCTATGGCTCCCACTGCTATGCCTTGTTTTTGTCACCAAAA 219

Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              :::||||||||||||:::      |||   |||:::   :::   |||:::|||||||||
Db        220 TCCTGGACAGATGCAGATCTGGCCTGCCAGAAGCGGCCCTCT---GGAAACCTGGTGTCT 276

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              :::   |||       |||       ||||||      |||:::   :::   |||   :::
Db        277 GTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTG---AAGAGCATTGGTAAC 333

Qy         62 AlaLysIleHisValTrpIleGlyLeuArg------AlaGlnAsnLysGluLysGlnCys 79
              :::      :::||||||||||||||||          |||      |||
Db        334 AGCTACTCATACGTCTGGATTGGGCTCCATGACCCCACACAGGGCACCGAGCCCAATGGA 393

Qy         80 ---SerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGlu 98
                 |||||||||           ::::::         |||
Db        394 GAAGGTTGGGAGTGGAGTAGCAGTGATGTGATGAATTACTTTGCATGGGAGAGAAATCCC 453

Qy         99 SerLys--------------LysCysLeuGlyValHisIleGluThrGlyPheHisLys 113
              |||                |||      :::         |||   |||   :::
Db        454 TCCACCATCTCAAGCCCCGGCCACTGTGCGAGCCTGTCGAGAAGCACAGCATTTCTGAGG 513

Qy        114 TrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||::::::::   |||      :::   |||:::|||||:::
Db        514 TGGAAAGATTATAACTGTAATGTGAGGTTACCCTATGTCTGCAAG 558


Search completed: December  8, 2003, 17:27:00
Job time : 106.726 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 16:08:50 ; Search time 513.051 Seconds
(without alignments)
835.675 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091002_22963/app_query.fasta_1.725
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09938114_@CGN_1_1_504_@runat_08122003_091002_22963
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	653	91.3	544	11	US-09-938-114-1	Sequence 1, Appli
	2	477	66.7	721	10	US-09-929-230-4	Sequence 4, Appli
	3	477	66.7	721	13	US-10-226-420-4	Sequence 4, Appli
	4	434	60.7	580	10	US-09-929-230-7	Sequence 7, Appli
	5	434	60.7	580	13	US-10-226-420-7	Sequence 7, Appli
	6	397	55.5	456	10	US-09-929-230-6	Sequence 6, Appli
	7	397	55.5	456	13	US-10-226-420-6	Sequence 6, Appli
	8	387	54.1	725	10	US-09-929-230-10	Sequence 10, Appl
	9	387	54.1	725	13	US-10-226-420-10	Sequence 10, Appl
	10	378	52.9	432	10	US-09-929-230-9	Sequence 9, Appli
	11	378	52.9	432	13	US-10-226-420-9	Sequence 9, Appli
	12	329	46.0	474	10	US-09-929-230-12	Sequence 12, Appl
	13	329	46.0	474	13	US-10-226-420-12	Sequence 12, Appl
	14	286.5	40.1	690	10	US-09-969-763-2	Sequence 2, Appli
	15	264	36.9	483	10	US-09-929-230-1	Sequence 1, Appli
	16	264	36.9	483	13	US-10-226-420-1	Sequence 1, Appli
	17	210	29.4	453	10	US-09-929-230-3	Sequence 3, Appli
	18	210	29.4	453	13	US-10-226-420-3	Sequence 3, Appli
	19	201.5	28.2	562	9	US-09-764-870-174	Sequence 174, App
	20	201.5	28.2	562	15	US-10-125-540-174	Sequence 174, App
	21	201.5	28.2	777	9	US-09-739-262-1	Sequence 1, Appli
	22	201.5	28.2	777	13	US-10-133-937-85	Sequence 85, Appl
	23	201.5	28.2	800	13	US-09-997-003-11	Sequence 11, Appl
	24	201.5	28.2	843	9	US-09-925-301-340	Sequence 340, App
	25	201.5	28.2	843	13	US-09-997-003-24	Sequence 24, Appl
	26	201.5	28.2	1066	9	US-09-764-870-15	Sequence 15, Appl
	27	201.5	28.2	1066	15	US-10-125-540-15	Sequence 15, Appl
	28	201.5	28.2	5641	15	US-10-015-219-1733	Sequence 1733, Ap
	29	201.5	28.2	5709	15	US-10-015-219-1734	Sequence 1734, Ap
	30	201.5	28.2	6214	15	US-10-015-219-1737	Sequence 1737, Ap
c	31	193.5	27.1	710	15	US-10-060-036-498	Sequence 498, App
c	32	192.5	26.9	539	15	US-10-060-036-3668	Sequence 3668, Ap
c	33	192.5	26.9	541	15	US-10-060-036-4035	Sequence 4035, Ap
c	34	192.5	26.9	554	15	US-10-060-036-3362	Sequence 3362, Ap
c	35	192.5	26.9	567	15	US-10-060-036-3129	Sequence 3129, Ap
	36	192.5	26.9	637	15	US-10-060-036-3637	Sequence 3637, Ap
c	37	192.5	26.9	644	15	US-10-060-036-3764	Sequence 3764, Ap
	38	192.5	26.9	767	10	US-09-920-300A-1748	Sequence 1748, Ap
	39	192.5	26.9	767	13	US-10-099-926-1748	Sequence 1748, Ap
	40	192.5	26.9	767	14	US-10-033-528-1748	Sequence 1748, Ap
	41	192.5	26.9	845	9	US-09-925-297-294	Sequence 294, App
c	42	189.5	26.5	548	15	US-10-060-036-3709	Sequence 3709, Ap
c	43	188.5	26.4	534	15	US-10-060-036-3517	Sequence 3517, Ap
	44	188.5	26.4	642	10	US-09-920-300A-1759	Sequence 1759, Ap
	45	188.5	26.4	642	13	US-10-099-926-1759	Sequence 1759, Ap

ALIGNMENTS

```
RESULT 1
US-09-938-114-1
; Sequence 1, Application US/09938114
; Publication No. US20030022350A1
;  GENERAL INFORMATION:
;       APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                  Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                  Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                  Jin-Guo Ding, Fang Rong, Yan Liu and
;                  Hui-Ran Chen
;       TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                           VENOM OF AGKISTRODON ACUTUS
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
```

```
;              COMPUTER: IBM Compatible
;              OPERATING SYSTEM: IBM P.C. DOS 5.0
;              SOFTWARE: FastSEQ for Windows 2.0
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/938,114
;              FILING DATE: 23-Aug-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/058,740
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Chen, Anthony C.
;              REGISTRATION NUMBER: 38,673
;              REFERENCE/DOCKET NUMBER: 233/298
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (213) 489-1600
;              TELEFAX: (213) 955-0440
;              TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 1:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 544 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-938-114-1

Alignment Scores:
Pred. No.:                2.44e-82        Length:          544
Score:                    653.00          Matches:         119
Percent Similarity:       99.17%          Conservative:    0
Best Local Similarity:    99.17%          Mismatches:      1
Query Match:              91.33%          Indels:          0
DB:                       11              Gaps:            0

US-09-938-114-2 (1-129) x US-09-938-114-1 (1-544)

Qy          10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db           1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

Qy          30 CysThrLysGlnValAsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAsp 49
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 TGCACGAAGCAGGTGAACGGGGGGCATCTGGTCTCTATCGAAAGCTCCGGAGAAGCAGAC 120

Qy          50 PheValGlyGlnLeuIleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGly 69
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TTTGTGGGCCAGTTGATTGCTCAGAAGATAAAGTCAGCCAAAATCCATGTCTGGATCGGA 180

Qy          70 LeuArgAlaGlnAsnLysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIle 89
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CTGAGGGCTCAAAACAAAGAAAAGCAATGCAGCATAGAGTGGAGCGATGGCTCCAGCATC 240

Qy          90 SerLysGluAsnTrpIleGluGluGluSerLysLysCysLeuGlyValHisIleGluThr 109
               |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AGTTATGAGAATTGGATTGAAGAAGAATCCAAAAAGTGTCTTGGGGTGCACATAGAGACA 300

Qy         110 GlyPheHisLysTrpGluAsnPheTyrCysGluGlnGlnAspProPheValCysGluAla 129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGGTTTCATAAGTGGGAGAATTTTTACTGTGAACAACAAGATCCTTTTGTCTGCGAGGCA 360

RESULT 2
US-09-929-230-4
; Sequence 4, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;    LENGTH: 721
;    TYPE: DNA
;    ORGANISM: Sistrurus miliarius
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (91)...(546)
;    NAME/KEY: misc_feature
;    LOCATION: (0)...(0)
;    OTHER INFORMATION: Zsnk3
US-09-929-230-4

Alignment Scores:
Pred. No.:                2.41e-57        Length:          721
Score:                    477.00          Matches:         83
Percent Similarity:       76.74%          Conservative:    16
Best Local Similarity:    64.34%          Mismatches:      30
Query Match:              66.71%          Indels:          0
DB:                       10              Gaps:            0

US-09-938-114-2 (1-129) x US-09-929-230-4 (1-721)

Qy           1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
               ||||||   |||   ||||||||||||:::   |||||||||:::||||||||||||
Db         160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy          21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
               |||||||||   |||||||||   ||||||:::::|||      ||||||||||||||||
Db         220 AAGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTC 279

Qy          41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
               ||||||||||||||||   ||||||   ||||||   ||||||:::   :::      :::
Db         280 TCTATCGAAAGCTCCGAAGAAGCAGCCTTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGG 339

Qy          61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
               |||   :::::::::|||||||||||||||   |||   ||||||||||||||||||
Db         340 AGAGCCATTCTCTATATCTGGATCGGACTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGC 399

Qy          81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
               :::|||||||||||||||||:::|||   |||||||||||||||   |||||||||
Db         400 GCGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 459

Qy         101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
               |||||||||:::          |||   :::||||||||   |||   ||||||
Db         460 ACATGTCTTGGGCTGCAACAAGGCACAAATTATCATAAGTGGGTCAATATTTACTGTGGA 519

Qy         121 GlnGlnAspProPheValCysGluAla 129
               :::   :::||||||||||||||||||
Db         520 GAAATAAATCCTTTTGTCTGCGAGGCA 546

RESULT 3
US-10-226-420-4
; Sequence 4, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;    LENGTH: 721
;    TYPE: DNA
;    ORGANISM: Sistrurus miliarius
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (91)...(546)
;    FEATURE:
;    NAME/KEY: misc_feature
```

```
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk3
US-10-226-420-4

Alignment Scores:
Pred. No.:              2.41e-57        Length:         721
Score:                  477.00          Matches:        83
Percent Similarity:     76.74%          Conservative:   16
Best Local Similarity:  64.34%          Mismatches:     30
Query Match:            66.71%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-2 (1-129) x US-10-226-420-4 (1-721)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              |||||| ||| |||||||||||:::  |||||||||:::|||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||  |||||||||  ||||||:::::|||       ||||||||||||||
Db        220 AAGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGGAGGGCGGGCATCTCGTC 279

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||||  ||||||  ||||||  ||||||:::  :::      :::
Db        280 TCTATCGAAAGCTCCGAAGAAGCAGCCTTTGTGGCCCAGCTGGTCCCTGAGAACAGGAGG 339

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||  :::::::::|||||||||||||||  |||  ||||||||||||||||||
Db        340 AGAGCCATTCTCTATATCTGGATCGGACTGAGGGTTCAAGGCAAAGAGAAGCAATGCAGC 399

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||:::|||  |||||||||||||||  |||||||||
Db        400 GCGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGGATTGAAGCAGAATCCAAA 459

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              |||||||||:::        |||  :::|||||||||  |||  ||||||
Db        460 ACATGTCTTGGGCTGCAACAAGGCACAAATTATCATAAGTGGGTCAATATTTACTGTGGA 519

Qy        121 GlnGlnAspProPheValCysGluAla 129
              :::   :::||||||||||||||||||
Db        520 GAAATAAATCCTTTTGTCTGCGAGGCA 546

RESULT 4
US-09-929-230-7
; Sequence 7, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(434)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk4
US-09-929-230-7

Alignment Scores:
Pred. No.:              2.04e-51        Length:         580
Score:                  434.00          Matches:        75
Percent Similarity:     74.22%          Conservative:   20
Best Local Similarity:  58.59%          Mismatches:     33
Query Match:            60.70%          Indels:         0
DB:                     10              Gaps:           0

US-09-938-114-2 (1-129) x US-09-929-230-7 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||  |||||||||  :::|||:::  :::||||||:::|||  ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              :::||||||  |||||||||  ||||||:::::|||       ||||||||||||||
Db         96 AGGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTC 155

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||  |||||||||  ||||||  ||||||:::|||:::  ||||||
Db        156 TCTATTGAAAGCGACGGAGAAGCAGCCTTTGTGGCCCAGCTGGTCGCTGAGAACATCAAG 215

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||       |||||||||||||||||  |||  :::|||||||||||||||||
Db        216 CAAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGGCGAAGAGAAGCAATGCAGC 275

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||||||||||:::::  ||||||  |||:::      :::|||
Db        276 ACGAAGTGGAGCGATGGCTCCAGCGTCAATTATGAGAACCTGATTAAACATGCGACCAAA 335

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||  |||:::      ||||||||||||      |||  |||  :::|||
Db        336 AAGTGTTTTGGGCTGAAAAAAGAGACAGGGTTTCGCACGTGGCGCAATGTTCACTGTACA 395

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||:::  |||:::|||:::
Db        396 CAACAAAATCTTTTCATGTGCAAG 419

RESULT 5
US-10-226-420-7
; Sequence 7, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 580
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(434)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk4
US-10-226-420-7

Alignment Scores:
Pred. No.:              2.04e-51        Length:         580
Score:                  434.00          Matches:        75
Percent Similarity:     74.22%          Conservative:   20
Best Local Similarity:  58.59%          Mismatches:     33
Query Match:            60.70%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-2 (1-129) x US-10-226-420-7 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||  |||||||||  :::|||:::  :::||||||:::|||  ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
```

```
Db         96 AGGACGTGGGACGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCGTC 155

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        156 TCTATTGAAAGCGACGGAGAAGCAGCCTTTGTGGCCCAGCTGGTCGCTGAGAACATCAAG 215

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        216 CAAAACAAATATGATGTCTGGATCGGACTGAGGATTCAAGGCGAAGAGAAGCAATGCAGC 275

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        276 ACGAAGTGGAGCGATGGCTCCAGCGTCAATTATGAGAACCTGATTAAACATGCGACCAAA 335

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        336 AAGTGTTTTGGGCTGAAAAAAGAGACAGGGTTTCGCACGTGGCGCAATGTTCACTGTACA 395

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        396 CAACAAAATCTTTTCATGTGCAAG 419
```

RESULT 6
US-09-929-230-6
; Sequence 6, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
; NAME/KEY: misc_feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
; LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
; LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
; LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
; OTHER INFORMATION: n = A,T,C or G
US-09-929-230-6

Alignment Scores:

Pred. No.:	2.39e-46	Length:	456
Score:	397.00	Matches:	66
Percent Similarity:	59.69%	Conservative:	11
Best Local Similarity:	51.16%	Mismatches:	52
Query Match:	55.52%	Indels:	0
DB:	10	Gaps:	0

US-09-938-114-2 (1-129) x US-09-929-230-6 (1-456)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        130 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNGARGGNGGNCAYYTNGTN 189
```

```
Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        190 WSNATHGARWSNWSNGARGARGCNGCNTTYGTNGCNCARYTNGTNCCNGARAAYMGNMGN 249

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        250 MGNGCNATHYTNTAYATHTGGATHGGNYTNMGNGTNCARGGNAARGARAARCARTGYWSN 309

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        310 GCNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYTGGATHGARGCNGARWSNAAR 369

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        370 ACNTGYYTNGGNYTNCARCARGGNACNAAYTAYCAYAARTGGGTNAAYATHTAYTGYGGN 429

Qy        121 GlnGlnAspProPheValCysGluAla 129

Db        430 GARATHAAYCCNTTYGTNTGYGARGCN 456
```

RESULT 7
US-10-226-420-6
; Sequence 6, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
; LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
; LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
; LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
; OTHER INFORMATION: n = A,T,C or G
US-10-226-420-6

Alignment Scores:

Pred. No.:	2.39e-46	Length:	456
Score:	397.00	Matches:	66
Percent Similarity:	59.69%	Conservative:	11
Best Local Similarity:	51.16%	Mismatches:	52
Query Match:	55.52%	Indels:	0
DB:	13	Gaps:	0

US-09-938-114-2 (1-129) x US-10-226-420-6 (1-456)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        130 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNGARGGNGGNCAYYTNGTN 189

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
```

```
Db        190 WSNATHGARWSNWSNGARGARGCNGCNTTYGTNGCNCARYTNGTNCCNGARAAYMGNMGN 249

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer  80

Db        250 MGNGCNATHYTNTAYATHTGGATHGGNYTNMGNGTNCARGGNAARGARAARCARTGYWSN 309

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        310 GCNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYTGGATHGARGCNGARWSNAAR 369

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        370 ACNTGYYTNGGNYTNCARCARGGNACNAAYTAYCAYAARTGGGTNAAYATHTAYTGYGGN 429

Qy        121 GlnGlnAspProPheValCysGluAla 129

Db        430 GARATHAAYCCNTTYGTNTGYGARGCN 456
```

RESULT 8
US-09-929-230-10
; Sequence 10, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk5
US-09-929-230-10

Alignment Scores:

Pred. No.:	1.19e-44	Length:	725
Score:	387.00	Matches:	67
Percent Similarity:	71.88%	Conservative:	25
Best Local Similarity:	52.34%	Mismatches:	36
Query Match:	54.13%	Indels:	0
DB:	10	Gaps:	0

US-09-938-114-2 (1-129) x US-09-929-230-10 (1-725)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db        163 AATTGTCCCTCTGGTTGGTTCGCCTACGATCAGTATTGCTACAGGGTCATCAAACGACTC 222

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        223 AAGACCTGGGACGATGCAGAGCGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCG 282

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        283 TCTGTCGAAAACGATGAAGAAGCAGTCTTTCTGGCCCAGTTGGTCGCTGCGAACATAAAG 342

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        343 CAAAACCAATACTATGTCTGGATTGGACTGAGGATTCAAAACAAAGGACAGCAATGCAGC 402

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        403 ACGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACCTGGTTAAATCACATTCCAAA 462

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        463 AAGTGTTTTGGGCTGAAAAAAGAGACAGAGTTTCTTCAATGGTACAATACTGACTGCGAA 522

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        523 GAAAAAAACCTTTTCGTCTGCAAG 546
```

RESULT 9
US-10-226-420-10
; Sequence 10, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/10/226,420
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(561)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk5
US-10-226-420-10

Alignment Scores:

Pred. No.:	1.19e-44	Length:	725
Score:	387.00	Matches:	67
Percent Similarity:	71.88%	Conservative:	25
Best Local Similarity:	52.34%	Mismatches:	36
Query Match:	54.13%	Indels:	0
DB:	13	Gaps:	0

US-09-938-114-2 (1-129) x US-10-226-420-10 (1-725)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20

Db        163 AATTGTCCCTCTGGTTGGTTCGCCTACGATCAGTATTGCTACAGGGTCATCAAACGACTC 222

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40

Db        223 AAGACCTGGGACGATGCAGAGCGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTGGCG 282

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60

Db        283 TCTGTCGAAAACGATGAAGAAGCAGTCTTTCTGGCCCAGTTGGTCGCTGCGAACATAAAG 342

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80

Db        343 CAAAACCAATACTATGTCTGGATTGGACTGAGGATTCAAAACAAAGGACAGCAATGCAGC 402

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100

Db        403 ACGAAGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACCTGGTTAAATCACATTCCAAA 462

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120

Db        463 AAGTGTTTTGGGCTGAAAAAAGAGACAGAGTTTCTTCAATGGTACAATACTGACTGCGAA 522

Qy        121 GlnGlnAspProPheValCysGlu 128

Db        523 GAAAAAAACCTTTTCGTCTGCAAG 546
```

```
RESULT 10
US-09-929-230-9
; Sequence 9, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;   NAME/KEY: misc_feature
;   LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;   LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;   LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;   LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9

Alignment Scores:
Pred. No.:                1.06e-43        Length:          432
Score:                    378.00          Matches:         64
Percent Similarity:       62.50%          Conservative:    16
Best Local Similarity:    50.00%          Mismatches:      48
Query Match:              52.87%          Indels:          0
DB:                       10              Gaps:            0

US-09-938-114-2 (1-129) x US-09-929-230-9 (1-432)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||   ||||||||   ||||||   :::|||      |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGTN 153

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||      |||||||||   ||||||   |||   :::|||:::   ||||||
Db        154 WSNATHGARWSNGAYGGNGARGCNGCNTTYGTNGCNCARYTNGTNGCNGARAAYATHAAR 213

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||      ||||||||||||          |||   :::||||||||||||
Db        214 CARAAYAARTAYGAYGTNTGGATHGGNYTNMGNATHCARGGNGARGARAARCARTGYWSN 273

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      ::::::   ||||||   |||:::      :::|||
Db        274 ACNAARTGGWSNGAYGGNWSNWSNGTNAAYTAYGARAAYYTNATHAARCAYGCNACNAAR 333

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||          ||||||||||||      |||   |||   :::|||
Db        334 AARTGYTTYGGNYTNAARAARGARACNGGNTTYMGNACNTGGMGNAAYGTNCAYTGYACN 393

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||:::   |||:::|||:::
Db        394 CARCARAAYYTNTTYATGTGYAAR 417

RESULT 11
US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;   LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;   LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;   LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-9

Alignment Scores:
Pred. No.:                1.06e-43        Length:          432
Score:                    378.00          Matches:         64
Percent Similarity:       62.50%          Conservative:    16
Best Local Similarity:    50.00%          Mismatches:      48
Query Match:              52.87%          Indels:          0
DB:                       13              Gaps:            0

US-09-938-114-2 (1-129) x US-10-226-420-9 (1-432)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||   ||||||||   ||||||   :::|||      |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGTN 153

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||      |||||||||   ||||||   |||   :::|||:::   ||||||
Db        154 WSNATHGARWSNGAYGGNGARGCNGCNTTYGTNGCNCARYTNGTNGCNGARAAYATHAAR 213

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||      ||||||||||||          |||   :::||||||||||||
Db        214 CARAAYAARTAYGAYGTNTGGATHGGNYTNMGNATHCARGGNGARGARAARCARTGYWSN 273

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      ::::::   ||||||   |||:::      :::|||
Db        274 ACNAARTGGWSNGAYGGNWSNWSNGTNAAYTAYGARAAYYTNATHAARCAYGCNACNAAR 333

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||          ||||||||||||      |||   |||   :::|||
Db        334 AARTGYTTYGGNYTNAARAARGARACNGGNTTYMGNACNTGGMGNAAYGTNCAYTGYACN 393

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||:::   |||:::|||:::
Db        394 CARCARAAYYTNTTYATGTGYAAR 417

RESULT 12
US-09-929-230-12
; Sequence 12, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
```

```
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
;   LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
;   LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
;   LOCATION: 402, 411, 426, 447, 453, 465, 468
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-12

Alignment Scores:
Pred. No.:                9.81e-37      Length:          474
Score:                    329.00        Matches:         55
Percent Similarity:       58.59%        Conservative:    20
Best Local Similarity:    42.97%        Mismatches:      53
Query Match:              46.01%        Indels:          0
DB:                       10            Gaps:            0

US-09-938-114-2 (1-129) x US-09-929-230-12 (1-474)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||           |||   :::|||:::   :::||||||   |||   |||
Db         76 AAYTGYCCNWSNGGNTGGTTYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARMGNYTN 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||   ||||||||   ||||||   :::|||       |||||||||
Db        136 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGCN 195

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              :::|||:::      ||||||   |||       |||   :::|||       ||||||
Db        196 WSNGTNGARAAYGAYGARGARGCNGTNTTYYTNGCNCARYTNGTNGCNGCNAAYATHAAR 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 :::   :::||||||||||||          |||||||||   :::||||||
Db        256 CARAAYCARTAYTAYGTNTGGATHGGNYTNMGNATHCARAAYAARGGNCARCARTGYWSN 315

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      :::      ||||||   ::::::          |||
Db        316 ACNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYYTNGTNAARWSNCAYWSNAAR 375

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||          ||||||   |||   :::|||   |||       ||||||
Db        376 AARTGYTTYGGNYTNAARAARGARACNGARTTYYTNCARTGGTAYAAYACNGAYTGYGAR 435

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::::::::   |||||||||:::
Db        436 GARAARAAYYTNTTYGTNTGYAAR 459

RESULT 13
US-10-226-420-12
; Sequence 12, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:11.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 84, 87, 90, 99, 120, 123, 132, 135, 141, 153, 159,
;   LOCATION: 168, 177, 183, 186, 192, 195, 198, 201, 219, 222, 228, 231,
;   LOCATION: 237, 240, 243, 246, 273, 282, 285, 288, 303, 315, 318
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 327, 333, 336, 339, 342, 345, 357, 360, 366, 372, 387, 390,
;   LOCATION: 402, 411, 426, 447, 453, 465, 468
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-12

Alignment Scores:
Pred. No.:                9.81e-37      Length:          474
Score:                    329.00        Matches:         55
Percent Similarity:       58.59%        Conservative:    20
Best Local Similarity:    42.97%        Mismatches:      53
Query Match:              46.01%        Indels:          0
DB:                       13            Gaps:            0

US-09-938-114-2 (1-129) x US-10-226-420-12 (1-474)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||           |||   :::|||:::   :::||||||   |||   |||
Db         76 AAYTGYCCNWSNGGNTGGTTYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARMGNYTN 135

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||||   ||||||||   ||||||   :::|||       |||||||||
Db        136 AARACNTGGGAYGAYGCNGARMGNTTYTGYWSNGARCARGCNAARGGNGGNCAYYTNGCN 195

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              :::|||:::      ||||||   |||       |||   :::|||       ||||||
Db        196 WSNGTNGARAAYGAYGARGARGCNGTNTTYYTNGCNCARYTNGTNGCNGCNAAYATHAAR 255

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                 :::   :::||||||||||||          |||||||||   :::||||||
Db        256 CARAAYCARTAYTAYGTNTGGATHGGNYTNMGNATHCARAAYAARGGNCARCARTGYWSN 315

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||   ||||||      :::      ||||||   ::::::          |||
Db        316 ACNAARTGGWSNGAYGGNWSNWSNGTNWSNTAYGARAAYYTNGTNAARWSNCAYWSNAAR 375

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              ||||||   |||          ||||||   |||   :::|||   |||       ||||||
Db        376 AARTGYTTYGGNYTNAARAARGARACNGARTTYYTNCARTGGTAYAAYACNGAYTGYGAR 435

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::::::::   |||||||||:::
Db        436 GARAARAAYYTNTTYGTNTGYAAR 459

RESULT 14
US-09-969-763-2
; Sequence 2, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
```

```
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 690
;   TYPE: DNA
;   ORGANISM: Crotalus harridus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (66)..(512)
;   OTHER INFORMATION:
US-09-969-763-2

Alignment Scores:
Pred. No.:                1.65e-30        Length:         690
Score:                    286.50          Matches:        55
Percent Similarity:       60.16%          Conservative:   22
Best Local Similarity:    42.97%          Mismatches:     42
Query Match:              40.07%          Indels:         9
DB:                       10              Gaps:           2

US-09-938-114-2 (1-129) x US-09-969-763-2 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||||||||||:::   :::|||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              ||||||   ||||||||   ||||||::::::|||      ||||||||||:::
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGGCATCTCCTC 260

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||:::|||::::::   ||||||   ||||||      ::::::              :::
Db        261 TCTGTCGAAACCGCCCTAGAAGCATCCTTTGTGGACAATGTGCTCTATGCGAACAAAGAG 320

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                    ::::::|||||||||||||||   |||||||||   :::   |||
Db        321 TACCTCACACGTTATATCTGGATTGGACTGAGGGTTCAAAACAAAGGACAGCCATGC--- 377

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                           ||||||||||||   ||||||   :::      :::
Db        378 ------------------TCCAGCATCAGTTATGAGAACCTGGTT------GACCCATTT 413

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
              :::|||      |||      :::|||         :::|||                ||||||
Db        414 GAATGTTTTATGGTGAGCAGAGACACAAGGCTTCGTGAGTGGTTTAAAGTTGACTGTGAA 473

Qy        121 GlnGlnAspProPheValCysGlu 128
              ||||||      |||:::|||:::
Db        474 CAACAACATTCTTTCATATGCAAG 497

RESULT 15
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(455)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk2
US-09-929-230-1

Alignment Scores:
Pred. No.:                1.48e-27        Length:         483
Score:                    264.00          Matches:        55
Percent Similarity:       57.03%          Conservative:   18
Best Local Similarity:    42.97%          Mismatches:     49
Query Match:              36.92%          Indels:         6
DB:                       10              Gaps:           3

US-09-938-114-2 (1-129) x US-09-929-230-1 (1-483)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||:::   ||||||||||||||||||   :::
Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||||||||   ||||||||||||||||||      |||                |||
Db        120 AAAACCTGGGATGATGCAGAGAGTTTCTGCTACACACAGCACAGAGACAGCCGCCTGGCC 179

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||   ||||||   |||      |||||||||:::|||   :::|||   :::|||
Db        180 TCCATCCACAGCAGTGAAGAAGAAGCTTTTGTGGGCAAACTGGCCTCCCAAACTTTGAAA 239

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                        :::||||||||||||:::      :::      |||:::|||
Db        240 TTCACTTCC------ATGTGGATCGGACTGAAA------GATCTATGGAAAGAATGCAAA 287

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              :::|||||||||   :::   :::      :::   |||
Db        288 TGGCAGTGGAGCGATGACACCAAACTGGACTACAAAGCCTGGACTCGAAGA------CCC 341

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
               |||      :::   ::::::|||              |||   |||      ||||||
Db        342 TATTGTACAGTAATGGTAGTCAAGACAGATAGGATCTTTTGGTTCAATAGAGGTTGCGAA 401

Qy        121 GlnGlnAspProPheValCysGlu 128
              :::         |||||||||:::
Db        402 AAGACTGTATCTTTGTCTGCAAG 425

Search completed: December  8, 2003, 19:36:04
Job time : 516.051 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:54 ; Search time 25.8 Seconds
(without alignments)
480.843 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	685	95.8	152	2	JC7134	agkisacutacin alph
2	527	73.7	129	2	JC4329	coagulation factor
3	496	69.4	152	2	JC4690	coagulation factor
4	422.5	59.1	133	2	A47267	botrocetin alpha c
5	420.5	58.8	144	2	PC7027	aggretin alpha cha
6	345	48.3	131	2	JC5058	bitiscetin alpha c
7	291	40.7	146	2	JC4691	coagulation factor
8	280	39.2	146	2	JC7105	aggretin beta chai
9	274	38.3	146	2	JC7135	agkisacutacin beta
10	271	37.9	125	2	JC5059	bitiscetin beta ch
11	260	36.4	125	2	B47267	botrocetin beta ch
12	252	35.2	123	2	JC2415	echicetin beta cha
13	248	34.7	123	2	B42972	coagulation factor
14	231.5	32.4	135	2	A38609	lectin, galactose-
15	201.5	28.2	166	1	RGHU1A	regenerating islet
16	201.5	28.2	166	2	A45751	pancreatic stone p
17	199.5	27.9	1340	2	A39808	proteoglycan core
18	199.5	27.9	2327	2	T42630	aggrecan - bovine
19	197.5	27.6	1479	2	T42710	mannose receptor,
20	194	27.1	172	2	S32489	lectin - Iberian r
21	192.5	26.9	166	1	RGHU1B	regenerating islet
22	190.5	26.6	165	2	A47148	reg I, regeneratin
23	189.5	26.5	173	2	B47148	reg II, regenerati
24	189.5	26.5	2124	2	A28452	proteoglycan core
25	189.5	26.5	2132	1	A55182	aggrecan precursor
26	186.5	26.1	174	2	S54979	pancreatitis-assoc
27	186	26.0	175	2	A37194	pancreatic thread
28	186	26.0	175	2	S29822	pancreatitis-assoc
29	185.5	25.9	1268	2	S52781	neurocan - mouse
30	185	25.9	174	2	I83377	regenerating prote
31	184.5	25.8	1257	2	S28764	neurocan precursor
32	183.5	25.7	2109	1	I50421	aggrecan precursor
33	183.5	25.7	2415	1	A39086	aggrecan precursor
34	183	25.6	175	2	A49616	pancreatitis-assoc
35	182.5	25.5	165	2	A28351	pancreatic stone p
36	177	24.8	174	1	A48689	pancreatitis-assoc
37	171	23.9	178	2	T29536	hypothetical prote
38	171	23.9	1458	1	A49707	phospholipase A2 r
39	170	23.8	175	2	A41719	pancreatic stone p
40	168	23.5	1463	2	A53210	phospholipase A2 r
41	166.5	23.3	1456	1	A36563	mannose receptor p
42	166	23.2	1326	2	B56395	secretory phosphol
43	166	23.2	1465	2	A56395	secretory phosphol
44	163	22.8	331	1	LNMSER	IgE Fc receptor, l
45	162.5	22.7	3562	2	A47171	chondroitin sulfat

ALIGNMENTS

RESULT 1
JC7134
agkisacutacin alpha chain precursor - sharp-nosed viper
N;Alternate names: fibrinogenlytic venom protein
C;Species: Agkistrodon acutus (sharp-nosed viper)
C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000
C;Accession: JC7134; PC7037
R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.
Biochem. Biophys. Res. Commun. 265, 530-535, 1999
A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenlytic venom
A;Reference number: JC7134; MUID:20025379; PMID:10558903
A;Accession: JC7134
A;Molecule type: mRNA
A;Residues: 1-152 <CHE>
A;Cross-references: GB:AF176420
A;Experimental source: venom gland
A;Accession: PC7037
A;Molecule type: protein
A;Residues: 24-53;84-86;87-94;125-136;137-152 <CH2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: agkisacutacin alpha chain #status experimental <MAT>

Query Match 95.8%; Score 685; DB 2; Length 152;
Best Local Similarity 96.1%; Pred. No. 6e-59;
Matches 124; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              |||| |||||||||||||||||| |||||||||||||||||||||||||  |||||||
Db         24 DCSSGWSSYEGHCYKVFKQSKTWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         84 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISYENWIEEESKKCLGVHIETGFHKWENFYCE 143

Qy        121 QQDPFVCEA 129
              |||||||||
Db        144 QQDPFVCEA 152
```

RESULT 2
JC4329
coagulation factor IX-binding protein A chain - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 13-Mar-1998
C;Accession: JC4329
R;Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 118, 965-973, 1995
A;Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flav
A;Reference number: JC4329; MUID:96318509; PMID:8749314
A;Accession: JC4329

A;Molecule type: protein
A;Residues: 1-129 <ATO>
C;Comment: This protein binds calcium.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; calcium binding; venom
F;2-127/Domain: C-type lectin homology <LCH>
F;2-13,30-127,102-119/Disulfide bonds: #status predicted

Query Match 73.7%; Score 527; DB 2; Length 129;
Best Local Similarity 72.1%; Pred. No. 9e-44;
Matches 93; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| ||  ||| ||| |||:|  |||||||||:|||||| ||: : |:
Db          1 DCPSGWSSYEGHCYKPFKLYKTWDDAERFCTEQAKGGHLVSIESAGEADFVAQLVTENIQ 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              : | :|||||| | |||||| |||||||:| ||||| ||| |||:  |||| || | ||
Db         61 NTKSYVWIGLRVQGKEKQCSSEWSDGSSVSYENWIEAESKTCLGLEKETGFRKWVNIYCG 120

Qy        121 QQDPFVCEA 129
              ||:||||||
Db        121 QQNPFVCEA 129
```

RESULT 3
JC4690
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4690; A39332
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4690
A;Molecule type: mRNA
A;Residues: 1-152 <MAT1>
A;Cross-references: DDBJ:D83331; NID:g1402639; PIDN:BAA11887.1; PID:g1402640
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
otein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: A39332
A;Status: preliminary
A;Molecule type: protein
A;Residues: 24-152 <ATO>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; glycoprotein; hemolymph; lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: factor IX/X binding protein chain A #status predicted <MAT>
F;25-150/Domain: C-type lectin homology <LCH>
F;25-36,53-150,125-142/Disulfide bonds: #status predicted

Query Match 69.4%; Score 496; DB 2; Length 152;
Best Local Similarity 67.4%; Pred. No. 1.1e-40;
Matches 87; Conservative 11; Mismatches 31; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| |:: ||| |||  ||:|  | ||||||||||||| ||: | :|
Db         24 DCLSGWSSYEGHCYKAFEKYKTWEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| | | |||: |||||||:| ||||| ||| |||:  || | || | ||
Db         84 RLDFYIWIGLRVQGKVKQCNSEWSDGSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCG 143

Qy        121 QQDPFVCEA 129
              ||:||||||
Db        144 QQNPFVCEA 152
```

RESULT 4
A47267
botrocetin alpha chain - jararaca
N;Alternate names: two chain botrocetin alpha chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: A47267; B37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator pur
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: A47267
A;Molecule type: protein
A;Residues: 1-133 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124085)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: B37958
A;Molecule type: protein
A;Residues: 1-40 <FUJ>
C;Complex: heterodimer of alpha and beta (see PIR:B47267) chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-128/Domain: C-type lectin homology <LCH>
F;2-13,30-128,103-120/Disulfide bonds: #status experimental
F;80/Disulfide bonds: interchain (to beta-75) #status experimental

Query Match 59.1%; Score 422.5; DB 2; Length 133;
Best Local Similarity 56.9%; Pred. No. 1.1e-33;
Matches 74; Conservative 20; Mismatches 35; Indels 1; Gaps 1;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIE-SSGEADFVGQLIAQKI 59
              || | ||||||:||| |:|   | ||| ||::|  |||||||:  | | |||| |: : |
Db          1 DCPSGWSSYEGNCYKFFQQKMNWADAERFCSEQAKGGHLVSIKIYSKEKDFVGDLVTKNI 60

Qy         60 KSAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYC 119
              :|: :: ||||| :||||||| |||||||:| || :|   |||  :  : ||  | | ||
Db         61 QSSDLYAWIGLRVENKEKQCSSEWSDGSSVSYENVVERTVKKCFALEKDLGFVLWINLYC 120

Qy        120 EQQDPFVCEA 129
               |::||||::
Db        121 AQKNPFVCKS 130
```

RESULT 5
PC7027
aggretin alpha chain - Malayan pit viper (fragment)
C;Species: Calloselasma rhodostoma (Malayan pit viper)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C;Accession: PC7027
R;Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A;Title: Molecular cloning and sequence analysis of aggretin, a collagen-like platelet a
A;Reference number: PC7027; MUID:99443731; PMID:10512747
A;Accession: PC7027
A;Molecule type: mRNA
A;Residues: 1-144 <CHU>
A;Experimental source: venom gland
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; platelet aggregation; venom

Query Match 58.8%; Score 420.5; DB 2; Length 144;
Best Local Similarity 59.7%; Pred. No. 1.9e-33;
Matches 77; Conservative 16; Mismatches 35; Indels 1; Gaps 1;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||   || |: |||: | : ||| :|| ||  | || || ||||:||||||  ||:|| :
Db         12 DCDFGWSPYDQHCYQAFNEQKTWDEAEKFCRAQENGAHLASIESNGEADFVSWLISQKDE 71

Qy         61 SA-KIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYC 119
```

```
                  | : :|||||||||||:||| |||||||:| || |:  :|||  :   ||| || |:||
Db          72 LADEDYVWIGLRAQNKEQQCSSEWSDGSSVSYENLIDLHTKKCGALEKLTGFRKWVNYYC 131

Qy         120 EQQDPFVCE 128
               ||   |||:
Db         132 EQMHAFVCK 140
```

RESULT 6
JC5058
bitiscetin alpha chain - puff adder
N;Alternate names: von Willebrand factor modulator protein
C;Species: Bitis arietans (puff adder)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 12-Feb-1999
C;Accession: JC5058; JC5916
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
submitted to JIPID, January 1997
A;Description: Complete amino acid sequence of bitiscetin, a novel von willebrand factor
A;Reference number: JC5058
A;Contents: snake venom
A;Accession: JC5058
A;Molecule type: protein
A;Residues: 1-131 <MAT>
A;Experimental source: snake venom
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
Res. Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997
A;Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor modul
A;Reference number: JC5916
A;Accession: JC5916
A;Molecule type: protein
A;Residues: 1-131 <MA2>
A;Experimental source: venom
C;Comment: This protein is a modulator of a von Willebrand factor modulator.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: venom
F;4-125/Domain: C-type lectin homology <LCH>

  Query Match   48.3%; Score 345; DB 2; Length 131;
  Best Local Similarity   47.7%; Pred. No. 3.4e-26;
  Matches   61; Conservative   24; Mismatches   39; Indels   4; Gaps   2;

```
Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |  |||||:||||||||:  || ||| || :  | ||| ||:|  |||||  :| :| :
Db          4 CLPDWSSYKGHCYKVFKKVGTWEDAEKFCVE--NSGHLASIDSKEEADFVTKLASQTLTK 61

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
                  ||||| ::| :||| :|:||||:  ||    :|  || |: : | :  | :  | :
Db         62 FVYDAWIGLRDESKTQQCSPQWTDGSSVVYEN--VDEPTKCFGLDVHTEYRTWTDLPCGE 119

Qy        122 QDPFVCEA 129
              ::||:|::
Db        120 KNPFICKS 127
```

RESULT 7
JC4691
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4691; B39332; JC4330
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4691
A;Molecule type: mRNA
A;Residues: 1-146 <MAT1>
A;Cross-references: DDBJ:D83332; NID:g1402641; PIDN:BAA11888.1; PID:g1402642
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
otein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: B39332
A;Molecule type: protein
A;Residues: 24-146 <ATO>
R;Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 118, 965-973, 1995
A;Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flavo
A;Reference number: JC4329; MUID:96318509; PMID:8749314
A;Accession: JC4330
A;Molecule type: protein
A;Residues: 24-146 <AT2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; lectin; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-146/Product: factor IX/X binding protein chain B #status predicted <MAT>
F;25-142/Domain: C-type lectin homology <LCH>
F;25-36,53-142,119-134/Disulfide bonds: #status predicted

  Query Match   40.7%; Score 291; DB 2; Length 146;
  Best Local Similarity   46.1%; Pred. No. 6.3e-21;
  Matches   59; Conservative   18; Mismatches   43; Indels   8; Gaps   4;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||| | : | | |||:|||:|  |||||| :|| ||||| :|  |
Db         24 DCPSDWSSYEGHCYKPFSEPKNWADAENFCTQQHAGGHLVSFQSSEEADFVVKLAFQTFG 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                :    |:||   |   ||: :||: : :  : | ||   |  |: ::  :|| :  |
Db         84 HSIF--WMGL--SNVWNQCNWQWSNAAMLRYKAWAEE--SYC--VYFKSTNNKWRSRACR 135

Qy        121 QQDPFVCE 128
                 ||||
Db        136 MMAQFVCE 143
```

RESULT 8
JC7105
aggretin beta chain - Malayan pit viper
C;Species: Calloselasma rhodostoma (Malayan pit viper)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7105
R;Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A;Title: Molecular cloning and sequence analysis of aggretin, a collagen-like platelet ac
A;Reference number: PC7027; MUID:99443731; PMID:10512747
A;Accession: JC7105
A;Molecule type: mRNA
A;Residues: 1-146 <CHU>
A;Experimental source: venom gland
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; platelet aggregation; venom

  Query Match   39.2%; Score 280; DB 2; Length 146;
  Best Local Similarity   44.3%; Pred. No. 7.2e-20;
  Matches   58; Conservative   18; Mismatches   41; Indels   14; Gaps   5;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| | : | | ||| ||  |    |||| :|: ||||| :|   ::|
Db         24 DCPSGWSSYEGHCYKPFNEPKNWADAERFCKLQPKHSHLVSFQSAEEADFVVKLTRPRLK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCL---GVHIETGFHKWENF 117
              :    ||:||   |    |: :||||: :: ::|  :|  :||   ||| |     | |
Db         84 ANL--VWMGL--SNIWHGCNWQWSDGARLNYKDW--QEQSECLAFRGVHTE-----WLNM 132

Qy        118 YCEQQDPFVCE 128
               |     |||:
Db        133 DCSSTCSFVCK 143
```

RESULT 9
JC7135
agkisacutacin beta chain precursor - sharp-nosed viper
N;Alternate names: fibrinogenlytic venom protein

```
C;Species: Agkistrodon acutus (sharp-nosed viper)
C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000
C;Accession: JC7135; PC7038
R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.
Biochem. Biophys. Res. Commun. 265, 530-535, 1999
A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenlytic venom
A;Reference number: JC7134; MUID:20025379; PMID:10558903
A;Accession: JC7135
A;Molecule type: mRNA
A;Residues: 1-146 <CHE>
A;Cross-references: GB:AF176421
A;Experimental source: venom gland
A;Accession: PC7038
A;Molecule type: protein
A;Residues: 24-50;59-83;102-107;112-114 <CH2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-146/Product: agkisacutacin beta chain #status experimental <MAT>

  Query Match             38.3%;  Score 274;  DB 2;  Length 146;
  Best Local Similarity   45.3%;  Pred. No. 2.7e-19;
  Matches   58;  Conservative   16;  Mismatches   46;  Indels    8;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |:|||||||||| | : ||| ||| |||:|  | || |  || |||||  |    :|
Db         24 DCPSEWSSYEGHCYKPFDEPKTWADAEKFCTQQHKGSHLASFHSSEEADFVVTLTTPSLK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              :    |||||:  |    | :||||: :  ::|   |  :||     |  ::| :  |
Db         84 TDL--VWIGLK--NIWNGCYWKWSDGTKLDYKDW--REQFECL--VSRTVNNEWLSMDCG 135

Qy        121 QQDPFVCE 128
                 |||:
Db        136 TTCSFVCK 143


RESULT 10
JC5059
bitiscetin beta chain - puff adder
C;Species: Bitis arietans (puff adder)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 12-Feb-1999
C;Accession: JC5059; JC5917
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
submitted to JIPID, January 1997
A;Description: Complete amino acid sequence of bitiscetin, a novel von willebrand factor
A;Reference number: JC5058
A;Accession: JC5059
A;Molecule type: protein
A;Residues: 1-125 <MAT>
A;Experimental source: snake venom
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
Res. Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997
A;Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor modul
A;Reference number: JC5916
A;Accession: JC5917
A;Molecule type: protein
A;Residues: 1-125 <MA2>
A;Experimental source: venom
C;Comment: This protein is a modulator of a von Willebrand factor modulator.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: venom
F;4-121/Domain: C-type lectin homology <LCH>

  Query Match             37.9%;  Score 271;  DB 2;  Length 125;
  Best Local Similarity   41.4%;  Pred. No. 4.5e-19;
  Matches   53;  Conservative   23;  Mismatches   42;  Indels   10;  Gaps    5;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |  |||||:||||||||  ||| ||| || : ||||||:|: |  | :|: :|  :|::
Db          4 CLPDWSSYKGHCYKVFKVEKTWADAEKFCKELVNGGHLMSVNSREEGEFISKLALEKMRI 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFH-KWENFYCE 120
                : |||||   :  : | : |:||: :      :|    |    :   || ||  : |
Db         64 --VLVWIGL--SHFWRICPLRWTDGARLDYRALSDEPI--CF---VAESFHNKWIQWTCN 114

Qy        121 QQDPFVCE 128
                ::  |||:
Db        115 RKKSFVCK 122


RESULT 11
B47267
botrocetin beta chain - jararaca
N;Alternate names: two chain botrocetin beta chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: B47267; C37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator pu
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: B47267
A;Molecule type: protein
A;Residues: 1-125 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124086)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: C37958
A;Molecule type: protein
A;Residues: 1-40 <FUJ>
C;Complex: heterodimer of alpha (see PIR:A47267) and beta chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-121/Domain: C-type lectin homology <LCH>
F;2-13,30-121,98-113/Disulfide bonds: #status experimental
F;75/Disulfide bonds: interchain (to alpha-80) #status experimental

  Query Match             36.4%;  Score 260;  DB 2;  Length 125;
  Best Local Similarity   42.2%;  Pred. No. 5.2e-18;
  Matches   54;  Conservative   17;  Mismatches   51;  Indels    6;  Gaps    3;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||  |||||||||||: ||:   | ||| |||:|  | |||| :|  |||||  | :: :|
Db          1 DCPPDWSSYEGHCYRFFKEWMHWDDAEEFCTEQQTGAHLVSFQSKEEADFVRSLTSEMLK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                   |||||   :   :|  ||:||     :::       :|  |  :   :||    |
Db         61 GDV--VWIGL--SDVWNKCRFEWTDGMEFDYDDYYLIAEYEC--VASKPTNNKWWIIPCT 114

Qy        121 QQDPFVCE 128
                :   ||||
Db        115 RFKNFVCE 122


RESULT 12
JC2415
echicetin beta chain - saw-scaled viper
C;Species: Echis carinatus (saw-scaled viper)
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 15-Oct-1996
C;Accession: JC2415
R;Peng, M.; Holt, J.C.; Niewiarowski, S.
Biochem. Biophys. Res. Commun. 205, 68-72, 1994
A;Title: Isolation, characterization and amino acid sequence of echicetin beta subunit,
A;Reference number: JC2415; MUID:95091801; PMID:7999097
A;Accession: JC2415
A;Molecule type: protein
A;Residues: 1-123 <PEN>
A;Experimental source: venom
C;Comment: This protein inhibits agglutination of fixed platelets induced by several pla
llebrand factor and alboaggregins.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; dimer
```

```
F;2-119/Domain: C-type lectin homology <LCH>
F;2-13,30-119,96-111/Disulfide bonds: #status predicted

  Query Match             35.2%;  Score 252;  DB 2;  Length 123;
  Best Local Similarity   40.6%;  Pred. No. 3e-17;
  Matches   52;  Conservative   18;  Mismatches   50;  Indels    8;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|  ||| |||:||||||:   | ||| || ||| |||||  :| | ||:  |    :|
Db          1 NCLPDWSVYEGYCYKVFKERMNWADAEKFCMKQVKDGHLVSFRNSKEVDFMISLAFPMLK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                |||||   :  : |  |||||: :  : |  :  : |     :|  ::|    |
Db         61 MEL--VWIGL--SDYWRDCYWEWSDGAQLDYKAW--DNERHCFAA--KTTDNQWMRRKCS 112

Qy        121 QQDPFVCE 128
               :  |||:
Db        113 GEFYFVCK 120


RESULT 13
B42972
coagulation factor X activating enzyme (EC 3.4.24.-) light chain - Russell's viper
C;Species: Vipera russelli (Russell's viper)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C;Accession: B42972
R;Takeya, H.; Nishida, S.; Miyata, T.; Kawada, S.; Saisaka, Y.; Morita, T.; Iwanaga, S.
J. Biol. Chem. 267, 14109-14117, 1992
A;Title: Coagulation factor X activating enzyme from Russell's viper venom (RVV-X). A nc
A;Reference number: A42972; MUID:92332516; PMID:1629211
A;Contents: V. r. siamensis
A;Accession: B42972
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-123 <TAK>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:108408)
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hydrolase; metalloproteinase; venom; zinc
F;4-121/Domain: C-type lectin homology <LCH>
F;4-15,32-121,98-113/Disulfide bonds: #status predicted

  Query Match             34.7%;  Score 248;  DB 2;  Length 123;
  Best Local Similarity   43.8%;  Pred. No. 7.3e-17;
  Matches   56;  Conservative   16;  Mismatches   48;  Indels    8;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | | ||| |||| |   | ||||| |||:|  | ||||: |  | :||  ||:: ::
Db          3 DCPSGWLSYEQHCYKGFNDLKNWTDAEKFCTEQKKGSHLVSLHSREEEEFVVNLISENLE 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ||||   |  | | :|||  :: |   : |||  ||   : |   :|::  |
Db         63 YPA--TWIGL--GNMWKDCRMEWSDRGNV-KYKALAEES-YCL--IMITHEKEWKSMTCN 114

Qy        121 QQDPFVCE 128
                | ||:
Db        115 FIAPVVCK 122


RESULT 14
A38609
lectin, galactose-specific - western diamondback rattlesnake
C;Species: Crotalus atrox (western diamondback rattlesnake)
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 11-Aug-1995
C;Accession: A38609
R;Hirabayashi, J.; Kusunoki, T.; Kasai, K.
J. Biol. Chem. 266, 2320-2326, 1991
A;Title: Complete primary structure of a galactose-specific lectin from the venom of the
A;Reference number: A38609; MUID:91115849; PMID:1989986
A;Accession: A38609
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-135 <HIR>
C;Superfamily: tetranectin; C-type lectin homology
F;3-131/Domain: C-type lectin homology <LCH>
F;3-14,31-131,106-123/Disulfide bonds: #status predicted

  Query Match             32.4%;  Score 231.5;  DB 2;  Length 135;
  Best Local Similarity   36.6%;  Pred. No. 3.2e-15;
  Matches   49;  Conservative   21;  Mismatches   55;  Indels    9;  Gaps    3;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|  ||    | |||:| | ||| ||| || |   | || |    ||:  : : |:   |
Db          2 NCPLDWLPMNGLCYKIFNQLKTWEDAEMFCRKYKPGCHLASFHRYGESLEIAEYISDYHK 61

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIE------EESKKCLGVHIETGFHKW 114
               : :||||||  :|:|  | ||:| |      | :      :  : |: :   ||:  |
Db         62 GQE-NVWIGLR--DKKKDFSWEWTDRSCTDYLTWDKNQPDHYQNKEFCVELVSLTGYRLW 118

Qy        115 ENFYCEQQDPFVCE 128
               :  || :| |:|:
Db        119 NDQVCESKDAFLCQ 132


RESULT 15
RGHU1A
regenerating islet lectin 1-alpha precursor [validated] - human
N;Alternate names: lithostathine; pancreatic thread protein (PTP); reg I protein; reg1-al
N;Contains: pancreatic stone protein (PSP)
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 03-Aug-1995 #text_change 08-Dec-2000
C;Accession: A35197; B28351; S12950; S02767; S02419; S00113; S01471; A25246
R;Watanabe, T.; Yonekura, H.; Terazono, K.; Yamamoto, H.; Okamoto, H.
J. Biol. Chem. 265, 7432-7439, 1990
A;Title: Complete nucleotide sequence of human reg gene and its expression in normal and
 product of the gene.
A;Reference number: A35197; MUID:90237042; PMID:2332435
A;Accession: A35197
A;Molecule type: DNA
A;Residues: 1-166 <WAT>
A;Cross-references: GB:J05412
R;Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Okamotc
J. Biol. Chem. 263, 2111-2114, 1988
A;Title: A novel gene activated in regenerating islets.
A;Reference number: A92704; MUID:88115343; PMID:2963000
A;Accession: B28351
A;Molecule type: mRNA
A;Residues: 1-166 <TER>
A;Cross-references: GB:M18963; NID:g190978; PIDN:AAA36558.1; PID:g190979
R;Itoh, T.; Tsuzuki, H.; Katoh, T.; Teraoka, H.; Matsumoto, K.; Yoshida, N.; Terazono, K
FEBS Lett. 272, 85-88, 1990
A;Title: Isolation and characterization of human reg protein produced in Saccharomyces ce
A;Reference number: S12950; MUID:91032149; PMID:2226837
A;Accession: S12950
A;Molecule type: protein
A;Residues: 23-52;160-166 <ITO>
A;Note: sequence determined from protein isolated after human cDNA sequence was cloned ar
R;de Caro, A.M.; Adrich, Z.; Fournet, B.; Capon, C.; Bonicel, J.J.; de Caro, J.D.; Rovery
Biochim. Biophys. Acta 994, 281-284, 1989
A;Title: N-terminal sequence extension in the glycosylated forms of human pancreatic stor
A;Reference number: S02767; MUID:89150292; PMID:2493268
A;Accession: S02767
A;Molecule type: protein
A;Residues: 23-47 <DEC>
R;Rouimi, P.; de Caro, J.; Bonicel, J.; Rovery, M.; de Caro, A.
FEBS Lett. 229, 171-174, 1988
A;Title: The disulfide bridges of the immunoreactive forms of human pancreatic stone prot
A;Reference number: S02419; MUID:88152214; PMID:3345835
A;Accession: S02419
A;Molecule type: protein
A;Residues: 63-72;125-139;150-157;160-166 <ROU>
A;Note: disulfide bonds
R;de Caro, A.M.; Bonicel, J.J.; Rouimi, P.; de Caro, J.D.; Sarles, H.; Rovery, M.
Eur. J. Biochem. 168, 201-207, 1987
A;Title: Complete amino acid sequence of an immunoreactive form of human pancreatic stone
A;Reference number: S00113; MUID:88029417; PMID:3665916
```

A;Accession: S00113
A;Molecule type: protein
A;Residues: 34-166 <DE1>
R;Rouimi, P.; Bonicel, J.; Rovery, M.; de Caro, A.
FEBS Lett. 216, 195-199, 1987
A;Title: Cleavage of the Arg-Ile bond in the native polypeptide chain of human pancreati
A;Reference number: S01471; MUID:87219142; PMID:3108036
A;Accession: S01471
A;Molecule type: protein
A;Residues: 33-48 <RO2>
R;Montalto, G.; Bonicel, J.; Multigner, L.; Rovery, M.; Sarles, H.; De Caro, A.
Biochem. J. 238, 227-232, 1986
A;Title: Partial amino acid sequence of human pancreatic stone protein, a novel pancreat
A;Reference number: A25246; MUID:87099950; PMID:3541906
A;Accession: A25246
A;Molecule type: protein
A;Residues: 34-73,'X',75-87,'R',89-98 <MON>
C;Comment: This protein is found in pancreatic calculi of mammals. The tryptic-like clea
C;Comment: Intact regenerating islet lectin 1-alpha, lithostathine, inhibits the growth
C;Genetics:
A;Gene: GDB:REG1A; REG
A;Cross-references: GDB:132455; OMIM:167770
A;Map position: 2p12-2p12
A;Introns: 22/1; 61/3; 107/3; 145/1
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: glycoprotein; lectin; pancreas; pyroglutamic acid
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-166/Product: regenerating islet lectin 1-alpha #status experimental <MAT>
F;34-166/Product: pancreatic stone protein #status experimental <MAT2>
F;36-162/Domain: C-type lectin homology <LCH>
F;23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;27/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;33-34/Cleavage site: Arg-Ile (trypsin) #status experimental
F;36-47,64-162,137-154/Disulfide bonds: #status experimental

Query Match 28.2%; Score 201.5; DB 1; Length 166;
Best Local Similarity 33.3%; Pred. No. 3.2e-12;
Matches 44; Conservative 26; Mismatches 53; Indels 9; Gaps 5;

```
Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |    ::|  :||   :  :|| ||: :| : :| |:|||: :  |  ||  || :   :
Db         36 CPEGTNAYRSYCYYFNEDRETWVDADLYC-QNMNSGNLVSVLTQAEGAFVASLIKES-GT 93

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENW-IEEESK----KCLGVHIETGFHKWEN 116
                :|||||    | ::    || || :| ::| |   |       |: :   ||| ||::
Db         94 DDFNVWIGLHDPKKNRRW--HWSSGSLVSYKSWGIGAPSSVNPGYCVSLTSSTGFQKWKD 151

Qy        117 FYCEQQDPFVCE 128
               || :  |||:
Db        152 VPCEDKFSFVCK 163
```

Search completed: December 8, 2003, 09:55:34
Job time : 25.8 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:46:14 ; Search time 16.9543 Seconds
(without alignments)
357.812 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	496	69.4	152	1	IXA_TRIFL	P23806 trimeresuru
2	425	59.4	131	1	ABA1_TRIAB	P81111 trimeresuru
3	422.5	59.1	133	1	BOTA_BOTJA	P22029 bothrops ja
4	408	57.1	134	1	ABA2_TRIAB	P81112 trimeresuru
5	388	54.3	132	1	ABBA_TRIAB	P81115 trimeresuru
6	364	50.9	158	1	CVXA_CRODU	O93426 crotalus du
7	309	43.2	133	1	ECHA_ECHCA	P81017 echis carin
8	302	42.2	133	1	RHCA_AGKRH	P81397 agkistrodon
9	299	41.8	123	1	ABA4_TRIAB	P81114 trimeresuru
10	291	40.7	146	1	IXB_TRIFL	P23807 trimeresuru
11	273.5	38.3	127	1	CHBA_CROHO	P81508 crotalus ho
12	269	37.6	129	1	RHCB_AGKRH	P81398 agkistrodon
13	264	36.9	148	1	CVXB_CRODU	O93427 crotalus du
14	260	36.4	125	1	BOTB_BOTJA	P22030 bothrops ja
15	259	36.2	117	1	CHBB_CROHO	P81509 crotalus ho
16	252	35.2	123	1	ECHB_ECHCA	P81996 echis carin
17	235	32.9	118	1	ABBB_TRIAB	P81116 trimeresuru
18	231.5	32.4	135	1	LECG_CROAT	P21963 crotalus at
19	230	32.2	125	1	ABA3_TRIAB	P81113 trimeresuru
20	215.5	30.1	158	1	LECG_TRIST	Q9ygp1 trimeresuru
21	209	29.2	175	1	PAP2_MOUSE	O09037 mus musculu
22	201.5	28.2	166	1	LITA_HUMAN	P05451 homo sapien
23	199.5	27.9	2364	1	PGCA_BOVIN	P13608 bos taurus
24	199	27.8	132	1	ACAL_ANSAN	P83300 anser anser
25	194.5	27.2	174	1	PAP3_MOUSE	O09049 mus musculu
26	194	27.1	172	1	LECA_PLEWA	Q02988 pleurodeles
27	192.5	26.9	166	1	LITB_HUMAN	P48304 homo sapien
28	190.5	26.6	165	1	LIT1_MOUSE	P43137 mus musculu
29	189.5	26.5	173	1	LIT2_MOUSE	Q08731 mus musculu
30	189.5	26.5	2124	1	PGCA_RAT	P07897 rattus norv
31	189.5	26.5	2132	1	PGCA_MOUSE	Q61282 mus musculu
32	186.5	26.1	174	1	PAP3_RAT	P42854 rattus norv
33	186	26.0	175	1	LITH_BOVIN	P23132 bos taurus
34	186	26.0	175	1	PAP1_MOUSE	P35230 mus musculu
35	185.5	25.9	1268	1	PGCN_MOUSE	P55066 mus musculu
36	185.5	25.9	2333	1	PGCA_CANFA	Q28343 canis famil
37	184.5	25.8	1257	1	PGCN_RAT	P55067 rattus norv
38	184.5	25.8	1321	1	PGCN_HUMAN	O14594 homo sapien
39	183.5	25.7	2109	1	PGCA_CHICK	P07898 gallus gall
40	183.5	25.7	2415	1	PGCA_HUMAN	P16112 homo sapien
41	183	25.6	175	1	PAP1_HUMAN	Q06141 homo sapien
42	182.5	25.5	165	1	LITH_RAT	P10758 rattus norv
43	177	24.8	174	1	PAP2_RAT	P35231 rattus norv
44	175	24.5	175	1	PBCG_HUMAN	Q92778 homo sapien
45	171	23.9	1458	1	PA2R_RABIT	P49260 oryctolagus

ALIGNMENTS

RESULT 1
IXA_TRIFL
ID IXA_TRIFL STANDARD; PRT; 152 AA.
AC P23806; Q91246;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor IX/factor X-binding protein A chain precursor
DE (IX/X-BP).
OS Trimeresurus flavoviridis (Habu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=88087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96184662; PubMed=8645314;
RA Matsuzaki R., Yoshihara E., Yamada M., Shima K., Atoda H., Morita T.;
RT "cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
RT protein from snake venom.";
RL Biochem. Biophys. Res. Commun. 220:382-387(1996).
RN [2]
RP SEQUENCE OF 24-152.
RC TISSUE=Venom;
RX MEDLINE=91332000; PubMed=1831197;
RA Atoda H., Hyuga M., Morita T.;
RT "The primary structure of coagulation factor IX/factor X-binding
RT protein isolated from the venom of Trimeresurus flavoviridis.
RT Homology with asialoglycoprotein receptors, proteoglycan core
RT protein, tetranectin, and lymphocyte Fc epsilon receptor for
RT immunoglobulin E.";
RL J. Biol. Chem. 266:14903-14911(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=97331317; PubMed=9187649;
RA Mizuno H., Fujimoto Z., Koizumi M., Kano H., Atoda H., Morita T.;
RT "Structure of coagulation factors IX/X-binding protein, a heterodimer
RT of C-type lectin domains.";
RL Nat. Struct. Biol. 4:438-441(1997).
CC -!- FUNCTION: ANTICOAGULANT PROTEIN WHICH BINDS WITH FACTOR IX AND
CC FACTOR X IN THE PRESENCE OF CALCIUM WITH A 1 TO 1 STOICHIOMETRY.
CC -!- SUBUNIT: HETERODIMER OF CHAINS A AND B LINKED BY A DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC ----------

CC ----------
DR EMBL; D83331; BAA11887.1; -.
DR PIR; JC4690; JC4690.

```
DR   PDB; 1IXX; 06-MAY-98.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Calcium; Signal; 3D-structure.
FT   SIGNAL        1     23
FT   CHAIN        24    152       COAGULATION FACTOR IX/FACTOR X-BINDING
FT                                 PROTEIN A CHAIN.
FT   DOMAIN       24    152       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     25     36       BY SIMILARITY.
FT   DISULFID     53    150       BY SIMILARITY.
FT   DISULFID    102    102       INTERCHAIN (WITH C-98 OF B CHAIN).
FT   DISULFID    125    142       BY SIMILARITY.
FT   TURN         27     28
FT   STRAND       30     31
FT   STRAND       36     44
FT   HELIX        46     56
FT   TURN         58     59
FT   STRAND       61     62
FT   HELIX        68     81
FT   STRAND       89     95
FT   STRAND      105    105
FT   TURN        107    108
FT   STRAND      111    111
FT   STRAND      117    117
FT   HELIX       119    121
FT   STRAND      125    128
FT   HELIX       130    132
FT   TURN        133    134
FT   STRAND      137    140
FT   TURN        142    143
FT   STRAND      146    152
SQ   SEQUENCE   152 AA;  17213 MW;  FB3DDD2369009263 CRC64;

  Query Match             69.4%;  Score 496;  DB 1;  Length 152;
  Best Local Similarity   67.4%;  Pred. No. 2.8e-40;
  Matches   87;  Conservative   11;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| |:: ||| |||  ||:|  | ||||||||||||| ||: | :|
Db         24 DCLSGWSSYEGHCYKAFEKYKTWEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| | | |||: |||||||:| ||||| ||| |||:  || | || | ||
Db         84 RLDFYIWIGLRVQGKVKQCNSEWSDGSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCG 143

Qy        121 QQDPFVCEA 129
                ||:||||||
Db        144 QQNPFVCEA 152


RESULT 2
ABA1_TRIAB
ID   ABA1_TRIAB     STANDARD;      PRT;   131 AA.
AC   P81111;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alboaggregin A subunit 1.
OS   Trimeresurus albolabris (White-lipped pit viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=8765;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=98189535; PubMed=9531050;
RA   Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA   Calvete J.J., Niewiarowski S.;
RT   "Alboaggregins A and B. Structure and interaction with human
RT   platelets.";
RL   Thromb. Haemost. 79:609-613(1998).
CC   -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC       agglutination.
CC   -!- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4,
CC       disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        1    129       C-TYPE LECTIN (LONG FORM).
FT   DISULFID      2     13       BY SIMILARITY.
FT   DISULFID     30    127       BY SIMILARITY.
FT   DISULFID    102    119       BY SIMILARITY.
SQ   SEQUENCE   131 AA;  15427 MW;  B3569F5BF91F6624 CRC64;

  Query Match             59.4%;  Score 425;  DB 1;  Length 131;
  Best Local Similarity   57.0%;  Pred. No. 1.3e-33;
  Matches   73;  Conservative   22;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||: :||:||| :|| ||| ||::| | |||||||:|||||| ||::: |:
Db          1 DCPSDWSSYDQYCYRVFKRIQTWEDAERFCSEQANDGHLVSIESAGEADFVTQLVSENIR 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | | :|||||| | | :||| |||||||:  :|  | :::|| |:     | | | ||
Db         61 SEKHYVWIGLRVQGKGQQCSSEWSDGSSVHYDNLQENKTRKCYGLEKRAEFRTWSNVYCG 120

Qy        121 QQDPFVCE 128
                : ||||:
Db        121 HEYPFVCK 128


RESULT 3
BOTA_BOTJA
ID   BOTA_BOTJA     STANDARD;      PRT;   133 AA.
AC   P22029;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Botrocetin, alpha chain (Platelet coagglutinin).
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724;
RN   [1]
RP   SEQUENCE, AND DISULFIDE BONDS.
RC   TISSUE=Venom;
RX   MEDLINE=93157385; PubMed=8430107;
RA   Usami Y., Fujimura Y., Suzuki M., Ozeki Y., Nishio K., Fukui H.,
RA   Titani K.;
RT   "Primary structure of two-chain botrocetin, a von Willebrand factor
RT   modulator purified from the venom of Bothrops jararaca.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:928-932(1993).
RN   [2]
RP   SEQUENCE OF 1-40.
RC   TISSUE=Venom;
RX   MEDLINE=91129280; PubMed=1993206;
RA   Fujimura Y., Titani K., Usami Y., Suzuki M., Oyama R., Matsui T.,
RA   Fukui H., Sugimoto M., Ruggeri Z.M.;
RT   "Isolation and chemical characterization of two structurally and
RT   functionally distinct forms of botrocetin, the platelet coagglutinin
RT   isolated from the venom of Bothrops jararaca.";
RL   Biochemistry 30:1957-1964(1991).
RN   [3]
```

```
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX   MEDLINE=22118144; PubMed=12121649;
RA   Fukuda K., Doggett T.A., Bankston L.A., Cruz M.A., Diacovo T.G.,
RA   Liddington R.C.;
RT   "Structural basis of von Willebrand factor activation by the snake
RT   toxin botrocetin.";
RL   Structure 10:943-950(2002).
CC   -!- FUNCTION: Two-chain Botrocetin forms an activated complex with
CC       vWF, and the complex then binds to platelet GPIb, resulting in
CC       platelet agglutination.
CC   -!- FUNCTION: There are two distinct forms of the von Willebrand
CC       factor-dependent platelet coagglutinin. The dimeric form is
CC       34-times more active than the one-chain Botrocetin in promoting
CC       vWF binding to platelets.
CC   -!- SUBUNIT: Disulfide-linked dimer of an alpha and a beta chain.
CC       vWF and Botrocetin form a soluble complex.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   PIR; A47267; A47267.
DR   PDB; 1IJK; 17-JUL-02.
DR   PDB; 1FVU; 14-FEB-01.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; 3D-structure.
FT   DISULFID      2     13
FT   DISULFID     30    128
FT   DISULFID     80     80       INTERCHAIN (WITH C-75 IN BETA CHAIN).
FT   DISULFID    103    120
SQ   SEQUENCE   133 AA;  15215 MW;  E4CF4502946AC74B CRC64;

  Query Match             59.1%;  Score 422.5;  DB 1;  Length 133;
  Best Local Similarity   56.9%;  Pred. No. 2.3e-33;
  Matches   74;  Conservative   20;  Mismatches   35;  Indels    1;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIE-SSGEADFVGQLIAQKI 59
              || | ||||||:||| |:|   | ||| ||::|  |||||||:  | | |||| |: : |
Db          1 DCPSGWSSYEGNCYKFFQQKMNWADAERFCSEQAKGGHLVSIKIYSKEKDFVGDLVTKNI 60

Qy         60 KSAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYC 119
              :|: :: ||||| :||||||| |||||||:| || :|   |||  :  : ||  | | ||
Db         61 QSSDLYAWIGLRVENKEKQCSSEWSDGSSVSYENVVERTVKKCFALEKDLGFVLWINLYC 120

Qy        120 EQQDPFVCEA 129
               |::||||::
Db        121 AQKNPFVCKS 130


RESULT 4
ABA2_TRIAB
ID   ABA2_TRIAB     STANDARD;      PRT;   134 AA.
AC   P81112;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alboaggregin A subunit 2.
OS   Trimeresurus albolabris (White-lipped pit viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=8765;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=98189535; PubMed=9531050;
RA   Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA   Calvete J.J., Niewiarowski S.;
RT   "Alboaggregins A and B. Structure and interaction with human
RT   platelets.";
RL   Thromb. Haemost. 79:609-613(1998).
CC   -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC       agglutination.
CC   -!- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4,
CC       disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        3    131       C-TYPE LECTIN (LONG FORM).
FT   DISULFID      4     15       BY SIMILARITY.
FT   DISULFID     32    129       BY SIMILARITY.
FT   DISULFID    104    121       BY SIMILARITY.
SQ   SEQUENCE   134 AA;  15556 MW;  5F9D71FC86DE2435 CRC64;

  Query Match             57.1%;  Score 408;  DB 1;  Length 134;
  Best Local Similarity   56.7%;  Pred. No. 5.4e-32;
  Matches   72;  Conservative   21;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              |   ||:|: :||:|| : | | ||| || || : |||||||: |||||| |||:: |:|
Db          4 CLPGWSAYDQYCYRVFNEPKNWEDAERFCAKQADSGHLVSIETMGEADFVAQLISENIQS 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
               | :|||||: ||||:||| |||||||:: || |:   :||  :  |:|| || |  | |
Db         64 EKHYVWIGLKVQNKEQQCSSEWSDGSSVTYENLIKLYMRKCGALEQESGFRKWINLGCIQ 123

Qy        122 QDPFVCE 128
                :||||:
Db        124 LNPFVCK 130


RESULT 5
ABBA_TRIAB
ID   ABBA_TRIAB     STANDARD;      PRT;   132 AA.
AC   P81115;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alboaggregin B alpha subunit.
OS   Trimeresurus albolabris (White-lipped pit viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=8765;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=98189535; PubMed=9531050;
RA   Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA   Calvete J.J., Niewiarowski S.;
RT   "Alboaggregins A and B. Structure and interaction with human
RT   platelets.";
RL   Thromb. Haemost. 79:609-613(1998).
CC   -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC       agglutination.
CC   -!- SUBUNIT: Heterodimer of alpha and beta subunits, disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
```

```
FT   DOMAIN        1    129       C-TYPE LECTIN (LONG FORM).
FT   DISULFID      2     13       BY SIMILARITY.
FT   DISULFID     30    127       BY SIMILARITY.
FT   DISULFID     79     79       INTERCHAIN (WITH C-75 IN BETA CHAIN)
FT                                (BY SIMILARITY).
FT   DISULFID    102    119       BY SIMILARITY.
SQ   SEQUENCE   132 AA;  15419 MW;  6FAE64820383F16F CRC64;

  Query Match             54.3%;  Score 388;  DB 1;  Length 132;
  Best Local Similarity   52.3%;  Pred. No. 4.2e-30;
  Matches   67;  Conservative   21;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||:: :||:: |: ||| ||| ||::| | |||||||  || || :|::: :
Db          1 DCPSDWSSFKQYCYQIVKELKTWEDAEXFCSEQANDGHLVSIESYREAVFVAELLSENVX 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                | | |||| ||| :||| |||||||:: || |     ||  :  |: |  | | |||
Db         61 XXKYHXWIGLSVQNKXQQCSSEWSDGSSVTYENLIXXXXXKCFVLKKESEFRTWSNVYCE 120

Qy        121 QQDPFVCE 128
                |:  |:|:
Db        121 QKHIFMCK 128


RESULT 6
CVXA_CRODU
ID   CVXA_CRODU     STANDARD;      PRT;   158 AA.
AC   O93426;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Convulxin alpha precursor (CVX alpha).
OS   Crotalus durissus terrificus (South American rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8732;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-48; 61-68; 108-120; 139-145 AND
RP   149-153.
RC   TISSUE=Venom gland;
RX   MEDLINE=98324901; PubMed=9657980;
RA   Leduc M., Bon C.;
RT   "Cloning of subunits of convulxin, a collagen-like platelet-
RT   aggregating protein from Crotalus durissus terrificus venom.";
RL   Biochem. J. 333:389-393(1998).
CC   -!- FUNCTION: Binds to the platelet and collagen receptor,
CC       glycoprotein VI (GPVI).
CC   -!- SUBUNIT: Heterohexamer of three alpha chains and three beta
CC       chains; disulfide-linked.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y16348; CAA76181.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Glycoprotein; Signal.
FT   SIGNAL        1     23
FT   CHAIN        24    158       CONVULXIN ALPHA.
FT   DOMAIN       34    153       C-TYPE LECTIN.
FT   DISULFID     27     38       BY SIMILARITY.
FT   DISULFID     55    152       BY SIMILARITY.
FT   DISULFID    104    104       INTERCHAIN (WITH C-100 IN BETA CHAIN)
FT                                (POTENTIAL).
FT   DISULFID    127    144       BY SIMILARITY.
FT   DISULFID    158    158       INTERCHAIN (WITH C-26 IN BETA CHAIN)
FT                                (POTENTIAL).
SQ   SEQUENCE   158 AA;  18141 MW;  949F9C6D673E2318 CRC64;

  Query Match             50.9%;  Score 364;  DB 1;  Length 158;
  Best Local Similarity   52.8%;  Pred. No. 9.6e-28;
  Matches   67;  Conservative   18;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              | |||  |: |||::| :  | ||| |||||  | |||||:|: |||||  :: | |:
Db         27 CPSDWYYYDQHCYRIFNEEMNWEDAEWFCTKQAKGAHLVSIKSAKEADFVAWMVTQNIEE 86

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
              :  || |||| |||||||| :||||||:| :| ::    ||  :  |||| ||    |
Db         87 SFSHVSIGLRVQNKEKQCSTKWSDGSSVSYDNLLDLYITKCSLLKKETGFRKWFVASCIG 146

Qy        122 QDPFVCE 128
              : ||||:
Db        147 KIPFVCK 153


RESULT 7
ECHA_ECHCA
ID   ECHA_ECHCA     STANDARD;      PRT;   133 AA.
AC   P81017;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Echicetin alpha subunit.
OS   Echis carinatus (Saw-scaled viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Echis.
OX   NCBI_TaxID=40353;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=97250657; PubMed=9163349;
RA   Polgar J., Magnenat E.M., Peitsch M.C., Wells T.N.C., Saqi M.S.A.,
RA   Clemetson K.J.;
RT   "Amino acid sequence of the alpha subunit and computer modelling of
RT   the alpha and beta subunits of echicetin from the venom of Echis
RT   carinatus (saw-scaled viper).";
RL   Biochem. J. 323:533-537(1997).
RN   [2]
RP   CHARACTERIZATION.
RC   TISSUE=Venom;
RX   MEDLINE=93244424; PubMed=8481512;
RA   Peng M., Lu W., Beviglia L., Niewiarowski S., Kirby E.P.;
RT   "Echicetin: a snake venom protein that inhibits binding of von
RT   Willebrand factor and alboaggregins to platelet glycoprotein Ib.";
RL   Blood 81:2321-2328(1993).
CC   -!- FUNCTION: Binds to platelet GPIb and inhibits platelet
CC       agglutination.
CC   -!- SUBUNIT: Heterodimer of alpha and beta subunits.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DISULFID      4     15       BY SIMILARITY.
```

```
FT   DISULFID      31    127       BY SIMILARITY.
FT   DISULFID     102    119       BY SIMILARITY.
FT   DISULFID      81     81       INTERCHAIN (WITH C-75 IN BETA CHAIN) (BY
FT                                 SIMILARITY).
SQ   SEQUENCE   133 AA;  15803 MW;  8F11C2D0BDC70B16 CRC64;

  Query Match             43.2%;  Score 309;  DB 1;  Length 133;
  Best Local Similarity   47.3%;  Pred. No. 1.3e-22;
  Matches   62;  Conservative   24;  Mismatches   39;  Indels    6;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | || ||||||::|:  ||| :|| :|  | :|||||||||: :|:|| |||::|:
Db          3 DCLSGWSFYEGHCYQLFR-LKTWDEAEKYC-NQWDGGHLVSIESNAKAEFVAQLISRKLP 60

Qy         61 SAKI--HVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFY 118
               : |   |||||| ::| :||   |:| | :  |: :     ||  :  :| | ||
Db         61 KSAIEDRVWIGLRDRSKREQCGHLWTDNSFVHYEHVV--PPTKCFVLERQTEFRKWIAVN 118

Qy        119 CEQQDPFVCEA 129
              || : ||||:|
Db        119 CEFKFPFVCKA 129


RESULT 8
RHCA_AGKRH
ID   RHCA_AGKRH     STANDARD;      PRT;   133 AA.
AC   P81397;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Rhodocetin alpha subunit.
OS   Agkistrodon rhodostoma (Malayan pit viper) (Calloselasma rhodostoma).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Calloselasma.
OX   NCBI_TaxID=8717;
RN   [1]
RP   SEQUENCE, FUNCTION, SUBUNIT, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99303998; PubMed=10360956;
RA   Wang R., Kini R.M., Chung M.C.M.;
RT   "Rhodocetin, a novel platelet aggregation inhibitor from the venom of
RT   Calloselasma rhodostoma (Malayan pit viper): synergistic and
RT   noncovalent interaction between its subunits.";
RL   Biochemistry 38:7584-7593(1999).
CC   -!- FUNCTION: A potent inhibitor of collagen-induced platelet
CC       aggregation. Individually, neither subunit inhibits platelet
CC       aggregation. Both subunits are essential.
CC   -!- SUBUNIT: Heterodimer of one alpha and one beta subunit held
CC       together by noncovalent interactions rather than by intersubunit
CC       disulfide bridges.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=15955.90; MW_ERR=1.44; METHOD=Electrospray.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        1    129       C-TYPE LECTIN (LONG FORM).
FT   DISULFID      2     13       BY SIMILARITY.
FT   DISULFID     30    127       BY SIMILARITY.
FT   DISULFID    102    119       BY SIMILARITY.
SQ   SEQUENCE   133 AA;  15962 MW;  386EAC519DFC674D CRC64;

  Query Match             42.2%;  Score 302;  DB 1;  Length 133;
  Best Local Similarity   42.2%;  Pred. No. 6e-22;
  Matches   54;  Conservative   27;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||   ||| : :||: ||: ||| :|| |||:|    ||||:|:  || ||  ::    :

Db          1 DCPDGWSSTKSYCYRPFKEKKTWEEAERFCTEQEKEAHLVSMENRLEAVFVDMVMENNFE 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               :     ||||: :|| :: ::|||||||||| ||  |   :||  :  ::|  ||    ||
Db         61 NKIYRSWIGLKIENKGQRSNLEWSDGSSISYENLYEPYMEKCFLMDHQSGLPKWHTADCE 120

Qy        121 QQDPFVCE 128
               ::: |:|:
Db        121 EKNVFMCK 128


RESULT 9
ABA4_TRIAB
ID   ABA4_TRIAB     STANDARD;      PRT;   123 AA.
AC   P81114;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alboaggregin A subunit 4.
OS   Trimeresurus albolabris (White-lipped pit viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=8765;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=98189535; PubMed=9531050;
RA   Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RA   Calvete J.J., Niewiarowski S.;
RT   "Alboaggregins A and B. Structure and interaction with human
RT   platelets.";
RL   Thromb. Haemost. 79:609-613(1998).
CC   -!- FUNCTION: Binds to platelet GPIb/IX receptor system and stimulates
CC       agglutination.
CC   -!- SUBUNIT: Heterotetramer of the subunits 1, 2, 3 and 4,
CC       disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23807; 1IXX.
DR   InterPro; IPR002353; AntifreezeII.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR00356; ANTIFREEZEII.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        1    121       C-TYPE LECTIN (LONG FORM).
FT   DISULFID      2     13       BY SIMILARITY.
FT   DISULFID     30    119       BY SIMILARITY.
FT   DISULFID     96    111       BY SIMILARITY.
SQ   SEQUENCE   123 AA;  14365 MW;  D4CFBEE1219C9B1E CRC64;

  Query Match             41.8%;  Score 299;  DB 1;  Length 123;
  Best Local Similarity   47.7%;  Pred. No. 1.1e-21;
  Matches   61;  Conservative   18;  Mismatches   41;  Indels    8;  Gaps    5;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||:|| : : | ||| |||:|  | |||| :|| ||||| |:  : |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKFCTQQHKGSHLVSFQSSEEADFVVQM-TRPIL 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               :| : |||||   |   ||: :||||: :  : |   |  :||     |  ::| :  |
Db         60 NANL-VWIGL--SNLWNQCNSQWSDGTXLDYKXW--REQFECL--VSRTTNNEWLSMDCS 112

Qy        121 QQDPFVCE 128
                 ||||
Db        113 STHSFVCE 120


RESULT 10
IXB_TRIFL
```

```
ID   IXB_TRIFL      STANDARD;      PRT;   146 AA.
AC   P23807; Q91247;
DT   01-NOV-1991 (Rel. 20, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Coagulation factor IX/factor X-binding protein B chain precursor
DE   (IX/X-BP).
OS   Trimeresurus flavoviridis (Habu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=88087;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96184662; PubMed=8645314;
RA   Matsuzaki R., Yoshihara E., Yamada M., Shima K., Atoda H., Morita T.;
RT   "cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
RT   protein from snake venom.";
RL   Biochem. Biophys. Res. Commun. 220:382-387(1996).
RN   [2]
RP   SEQUENCE OF 24-146.
RC   TISSUE=Venom;
RX   MEDLINE=91332000; PubMed=1831197;
RA   Atoda H., Hyuga M., Morita T.;
RT   "The primary structure of coagulation factor IX/factor X-binding
RT   protein isolated from the venom of Trimeresurus flavoviridis.
RT   Homology with asialoglycoprotein receptors, proteoglycan core
RT   protein, tetranectin, and lymphocyte Fc epsilon receptor for
RT   immunoglobulin E.";
RL   J. Biol. Chem. 266:14903-14911(1991).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=97331317; PubMed=9187649;
RA   Mizuno H., Fujimoto Z., Koizumi M., Kano H., Atoda H., Morita T.;
RT   "Structure of coagulation factors IX/X-binding protein, a heterodimer
RT   of C-type lectin domains.";
RL   Nat. Struct. Biol. 4:438-441(1997).
CC   -!- FUNCTION: ANTICOAGULANT PROTEIN WHICH BINDS WITH FACTOR IX AND
CC       FACTOR X IN THE PRESENCE OF CALCIUM WITH A 1 TO 1 STOICHIOMETRY.
CC   -!- SUBUNIT: HETERODIMER OF CHAINS A AND B LINKED BY A DISULFIDE BOND.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D83332; BAA11888.1; -.
DR   PIR; JC4691; JC4691.
DR   PDB; 1IXX; 06-MAY-98.
DR   PDB; 1BJ3; 16-AUG-99.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Calcium; Signal; 3D-structure.
FT   SIGNAL        1     23
FT   CHAIN        24    146       COAGULATION FACTOR IX/FACTOR X-BINDING
FT                                 PROTEIN B CHAIN.
FT   DOMAIN       24    144       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     25     36       BY SIMILARITY.
FT   DISULFID     53    142       BY SIMILARITY.
FT   DISULFID     98     98       INTERCHAIN (WITH C-102 OF A CHAIN).
FT   DISULFID    119    134       BY SIMILARITY.
FT   TURN         27     28
FT   STRAND       30     32
FT   TURN         33     34
FT   STRAND       35     44
FT   HELIX        46     54
FT   TURN         55     56
FT   TURN         58     59
FT   STRAND       61     62
FT   HELIX        68     82
FT   STRAND       86     88
FT   TURN         94     97
FT   STRAND      100    102
FT   TURN        103    104
FT   STRAND      113    113
FT   STRAND      118    123
FT   TURN        124    125
FT   STRAND      126    133
FT   TURN        134    135
FT   STRAND      138    146
SQ   SEQUENCE   146 AA;  16922 MW;  8E1961C59F96757C CRC64;

  Query Match             40.7%;  Score 291;  DB 1;  Length 146;
  Best Local Similarity   46.1%;  Pred. No. 7.4e-21;
  Matches   59;  Conservative   18;  Mismatches   43;  Indels    8;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||| | : | | |||:|||:|  |||||| :|| ||||| :|  |
Db         24 DCPSDWSSYEGHCYKPFSEPKNWADAENFCTQQHAGGHLVSFQSSEEADFVVKLAFQTFG 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                :    |:||   |   ||: :||: : :  : | ||    |  |: ::  :|| :  |
Db         84 HSIF--WMGL--SNVWNQCNWQWSNAAMLRYKAWAEE--SYC--VYFKSTNNKWRSRACR 135

Qy        121 QQDPFVCE 128
                ||||
Db        136 MMAQFVCE 143


RESULT 11
CHBA_CROHO
ID   CHBA_CROHO     STANDARD;      PRT;   127 AA.
AC   P81508;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   CHH-B alpha subunit.
OS   Crotalus horridus horridus (Timber rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8747;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=96420502; PubMed=8823201;
RA   Andrews R.K., Kroll M.H., Ward C.M., Rose J.W., Scarborough R.M.,
RA   Smith A.I., Lopez J.A., Berndt M.C.;
RT   "Binding of a novel 50-kilodalton alboaggregin from Trimeresurus
RT   albolabris and related viper venom proteins to the platelet membrane
RT   glycoprotein Ib-IX-V complex. Effect on platelet aggregation and
RT   glycoprotein Ib-mediated platelet activation.";
RL   Biochemistry 35:12629-12639(1996).
CC   -!- FUNCTION: Binds to platelet GPiB/IX receptor system, inhibits VWF
CC       binding, and stimulates agglutination.
CC   -!- SUBUNIT: Heterodimer of alpha and beta subunits; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN       11    121       C-TYPE LECTIN.
FT   DISULFID      4     15       BY SIMILARITY.
FT   DISULFID     32    120       BY SIMILARITY.
FT   DISULFID     81     81       INTERCHAIN (WITH C-92 IN BETA CHAIN)
FT                                (POTENTIAL).
```

```
FT   DISULFID     95    112       BY SIMILARITY.
SQ   SEQUENCE   127 AA;  15162 MW;  B5DA100D383E3547 CRC64;

  Query Match           38.3%;  Score 273.5;  DB 1;  Length 127;
  Best Local Similarity   41.4%;  Pred. No. 2.9e-19;
  Matches   53;  Conservative   21;  Mismatches   45;  Indels    9;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :| | |||| : ||| |||  || ||| ||::|  | ||:|:|:: || ||  ::    :
Db          3 ECPSGWSSYDRYCYKPFKQEMTWADAERFCSEQAKGRHLLSVETALEASFVDNVLYANKE 62

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| ||| : |       |||  || ::     |  |  :|   :|    ||
Db         63 YLTRYIWIGLRVQNKGQPC-------SSIYSENLVDPFD--CFMVSRDTRLREWFKVDCE 113

Qy        121 QQDPFVCE 128
              ||  |:|:
Db        114 QQHSFICK 121


RESULT 12
RHCB_AGKRH
ID   RHCB_AGKRH     STANDARD;      PRT;   129 AA.
AC   P81398;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Rhodocetin beta subunit.
OS   Agkistrodon rhodostoma (Malayan pit viper) (Calloselasma rhodostoma).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Calloselasma.
OX   NCBI_TaxID=8717;
RN   [1]
RP   SEQUENCE, FUNCTION, SUBUNIT, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99303998; PubMed=10360956;
RA   Wang R., Kini R.M., Chung M.C.M.;
RT   "Rhodocetin, a novel platelet aggregation inhibitor from the venom of
RT   Calloselasma rhodostoma (Malayan pit viper): synergistic and
RT   noncovalent interaction between its subunits.";
RL   Biochemistry 38:7584-7593(1999).
CC   -!- FUNCTION: A potent inhibitor of collagen-induced platelet
CC       aggregation. Individually, neither subunit inhibits platelet
CC       aggregation. Both subunits are essential.
CC   -!- SUBUNIT: Heterodimer of one alpha and one beta subunit held
CC       together by noncovalent interactions rather than by intersubunit
CC       disulfide bridges.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=15184.53; MW_ERR=2.74; METHOD=Electrospray.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN       3    125       C-TYPE LECTIN (LONG FORM).
FT   DISULFID     4     15       BY SIMILARITY.
FT   DISULFID    98    115       BY SIMILARITY.
SQ   SEQUENCE   129 AA;  15190 MW;  EF4B318FAAC807AE CRC64;

  Query Match           37.6%;  Score 269;  DB 1;  Length 129;
  Best Local Similarity   39.4%;  Pred. No. 7.9e-19;
  Matches   50;  Conservative   23;  Mismatches   48;  Indels    6;  Gaps    1;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              | : ||: : :||| ||: ||| :|| || ||   ||||| |: ||||:  :|
Db          4 CPTTWSASKLYCYKPFKEKKTWIEAERFCAKQAENGHLVSIGSAAEADFLDLVIVVNFDK 63

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
               :  | ||  :|      ::|::|:|:| ||  |   :|| |    |  ||    ||:

Db         64 QRYRAWTGLTERN------LKWTNGASVSYENLYEPYIRKCFVVQPWEGKSKWYKADCEE 117

Qy        122 QDPFVCE 128
              :: |:|:
Db        118 KNAFLCK 124


RESULT 13
CVXB_CRODU
ID   CVXB_CRODU     STANDARD;      PRT;   148 AA.
AC   O93427;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Convulxin beta precursor (CVX beta).
OS   Crotalus durissus terrificus (South American rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8732;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 24-53; 99-109 AND 126-145.
RC   TISSUE=Venom gland;
RX   MEDLINE=98324901; PubMed=9657980;
RA   Leduc M., Bon C.;
RT   "Cloning of subunits of convulxin, a collagen-like platelet-
RT   aggregating protein from Crotalus durissus terrificus venom.";
RL   Biochem. J. 333:389-393(1998).
CC   -!- FUNCTION: Binds to the platelet and collagen receptor,
CC       glycoprotein VI (GPVI).
CC   -!- SUBUNIT: Heterohexamer of three alpha chains and three beta
CC       chains; disulfide-linked.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y16349; CAA76182.1; -.
DR   HSSP; P23807; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; Glycoprotein; Signal.
FT   SIGNAL       1     23
FT   CHAIN       24    148       CONVULXIN BETA.
FT   DOMAIN      34    145       C-TYPE LECTIN.
FT   DISULFID    26     26       INTERCHAIN (WITH C-158 IN ALPHA CHAIN)
FT                               (POTENTIAL).
FT   DISULFID    27     38       BY SIMILARITY.
FT   DISULFID    55    144       BY SIMILARITY.
FT   DISULFID   100    100       INTERCHAIN (WITH C-104 IN ALPHA CHAIN)
FT                               (POTENTIAL).
FT   DISULFID   121    136       BY SIMILARITY.
SQ   SEQUENCE   148 AA;  17402 MW;  94D7E3E1BC693B9F CRC64;

  Query Match           36.9%;  Score 264;  DB 1;  Length 148;
  Best Local Similarity   44.5%;  Pred. No. 2.7e-18;
  Matches   57;  Conservative   17;  Mismatches   44;  Indels   10;  Gaps    5;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              | | ||||: :|||||||  || ||| |||:|  | ||||  |: | ||| ::  | :||
Db         27 CPSHWSSYDRYCYKVFKQEMTWADAEKFCTQQHTGSHLVSFHSTEEVDFVVKMTHQSLKS 86

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGF-HKWENFYCE 120
                 |||  |  :|: :||||:    : | ||   :||   |   | ::| : |
Db         87 T--FFWIG--ANNIWNKCNWQWSDGTKPEYKEWHEE--FECL---ISRTFDNQWLSAPCS 137
```

```
Qy          121 QQDPFVCE 128
                |||:
Db          138 DTYSFVCK 145


RESULT 14
BOTB_BOTJA
ID   BOTB_BOTJA     STANDARD;      PRT;   125 AA.
AC   P22030;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Botrocetin, beta chain (Platelet coagglutinin).
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724;
RN   [1]
RP   SEQUENCE, AND DISULFIDE BONDS.
RC   TISSUE=Venom;
RX   MEDLINE=93157385; PubMed=8430107;
RA   Usami Y., Fujimura Y., Suzuki M., Ozeki Y., Nishio K., Fukui H.,
RA   Titani K.;
RT   "Primary structure of two-chain botrocetin, a von Willebrand factor
RT   modulator purified from the venom of Bothrops jararaca.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:928-932(1993).
RN   [2]
RP   SEQUENCE OF 1-40.
RC   TISSUE=Venom;
RX   MEDLINE=91129280; PubMed=1993206;
RA   Fujimura Y., Titani K., Usami Y., Suzuki M., Oyama R., Matsui T.,
RA   Fukui H., Sugimoto M., Ruggeri Z.M.;
RT   "Isolation and chemical characterization of two structurally and
RT   functionally distinct forms of botrocetin, the platelet coagglutinin
RT   isolated from the venom of Bothrops jararaca.";
RL   Biochemistry 30:1957-1964(1991).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX   MEDLINE=22118144; PubMed=12121649;
RA   Fukuda K., Doggett T.A., Bankston L.A., Cruz M.A., Diacovo T.G.,
RA   Liddington R.C.;
RT   "Structural basis of von Willebrand factor activation by the snake
RT   toxin botrocetin.";
RL   Structure 10:943-950(2002).
CC   -!- FUNCTION: Two-chain Botrocetin forms an activated complex with
CC       vWF, and the complex then binds to platelet GPIb, resulting in
CC       platelet agglutination.
CC   -!- FUNCTION: There are two distinct forms of the von Willebrand
CC       factor-dependent platelet coagglutinin. The dimeric form is
CC       34-times more active than the one-chain Botrocetin in promoting
CC       vWF binding to platelets.
CC   -!- SUBUNIT: Disulfide-linked dimer of an alpha and a beta chain.
CC       vWF and Botrocetin form a soluble complex.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   PIR; B47267; B47267.
DR   PDB; 1IJK; 17-JUL-02.
DR   PDB; 1FVU; 14-FEB-01.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin; 3D-structure.
FT   DISULFID      2     13
FT   DISULFID     30    121
FT   DISULFID     75     75       INTERCHAIN (WITH C-80 IN ALPHA CHAIN).
FT   DISULFID     98    113
SQ   SEQUENCE   125 AA;  15037 MW;  1ED2027ED817FCA0 CRC64;

  Query Match             36.4%;  Score 260;  DB 1;  Length 125;
  Best Local Similarity   42.2%;  Pred. No. 5.4e-18;
  Matches   54;  Conservative   17;  Mismatches   51;  Indels    6;  Gaps    3;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||  |||||||||||: ||:   | ||| |||:|  | |||| :|  ||||| | :: :|
Db          1 DCPPDWSSYEGHCYRFFKEWMHWDDAEEFCTEQQTGAHLVSFQSKEEADFVRSLTSEMLK 60

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                |||||   :   :|  ||:||     :::      :|  |  :   :||    |
Db         61 GDV--VWIGL--SDVWNKCRFEWTDGMEFDYDDYYLIAEYEC--VASKPTNNKWWIIPCT 114

Qy        121 QQDPFVCE 128
               :  ||||
Db        115 RFKNFVCE 122


RESULT 15
CHBB_CROHO
ID   CHBB_CROHO     STANDARD;      PRT;   117 AA.
AC   P81509;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   CHH-B beta subunit.
OS   Crotalus horridus horridus (Timber rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8747;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=96420502; PubMed=8823201;
RA   Andrews R.K., Kroll M.H., Ward C.M., Rose J.W., Scarborough R.M.,
RA   Smith A.I., Lopez J.A., Berndt M.C.;
RT   "Binding of a novel 50-kilodalton alboaggregin from Trimeresurus
RT   albolabris and related viper venom proteins to the platelet membrane
RT   glycoprotein Ib-IX-V complex. Effect on platelet aggregation and
RT   glycoprotein Ib-mediated platelet activation.";
RL   Biochemistry 35:12629-12639(1996).
CC   -!- FUNCTION: Binds to platelet GPiB/IX receptor system, inhibits VWF
CC       binding, and stimulates agglutination.
CC   -!- SUBUNIT: Heterodimer of alpha and beta subunits; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
FT   DOMAIN        9    116       C-TYPE LECTIN.
FT   DISULFID      2     13       BY SIMILARITY.
FT   DISULFID     30    115       BY SIMILARITY.
FT   DISULFID     92     92       INTERCHAIN (WITH C-81 IN ALPHA CHAIN)
FT                                (POTENTIAL).
FT   DISULFID     95    107       BY SIMILARITY.
SQ   SEQUENCE   117 AA;  13888 MW;  07835BBCB61E9EAD CRC64;

  Query Match             36.2%;  Score 259;  DB 1;  Length 117;
  Best Local Similarity   40.3%;  Pred. No. 6.3e-18;
  Matches   52;  Conservative   21;  Mismatches   42;  Indels   14;  Gaps    4;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || |||||||||||:||:|  || ||| |||:|  ||||||  || | ||:  ::   :
Db          1 DCPSDWSSYEGHCYRVFQQEMTWDDAEKFCTQQHTGGHLVSFRSSEEVDFLVSILKFDL- 59

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENW-IEEESKKCLGVHIETGFHKWENFYC 119
                |:| |     |::  ::||||: :: : |  | |   |          ::| :  |
Db         60 -----FWMGWRDIWNERR--LQWSDGTKVNYKAWSAEPECIVCRATD-----NQWLSTSC 107

Qy        120 EQQDPFVCE 128
               :   ||:
Db        108 SKTHNVVCK 116
```

Search completed: December 8, 2003, 09:53:09
Job time : 17.2876 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:24 ; Search time 62.6571 Seconds
(without alignments)
531.285 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	701	98.0	152	13	Q9IAM1	Q9iam1 agkistrodon
2	657	91.9	152	13	Q9DEF9	Q9def9 agkistrodon
3	657	91.9	152	13	Q8JIW0	Q8jiw0 agkistrodon
4	496	69.4	146	13	Q98UJ0	Q98uj0 trimeresuru
5	475	66.4	152	13	Q9DG39	Q9dg39 agkistrodon
6	460	64.3	154	13	Q9YGN5	Q9ygn5 agkistrodon
7	439	61.4	158	13	Q8UVC6	Q8uvc6 agkistrodon
8	439	61.4	158	13	Q8AYA5	Q8aya5 agkistrodon
9	420.5	58.8	136	13	Q9I841	Q9i841 agkistrodon
10	394	55.1	158	13	Q8JGT7	Q8jgt7 trimeresuru
11	391	54.7	158	13	Q8AV97	Q8av97 trimeresuru
12	390	54.5	131	13	Q9PSM9	Q9psm9 echis carin
13	382.5	53.5	142	13	Q9PSM6	Q9psm6 bothrops ja
14	380.5	53.2	155	13	Q8JIV6	Q8jiv6 agkistrodon
15	377.5	52.8	155	13	Q9DEA2	Q9dea2 agkistrodon
16	375	52.4	154	13	Q8JIV9	Q8jiv9 agkistrodon
17	355.5	49.7	157	13	Q9YGG9	Q9ygg9 agkistrodon
18	290	40.6	146	13	Q9DG31	Q9dg31 agkistrodon
19	290	40.6	146	13	Q8AYA4	Q8aya4 agkistrodon
20	289	40.4	146	13	Q8JIV7	Q8jiv7 agkistrodon
21	282	39.4	124	13	Q98SM5	Q98sm5 agkistrodon
22	280	39.2	146	13	Q9I840	Q9i840 agkistrodon
23	278	38.9	146	13	Q9IAM0	Q9iam0 agkistrodon
24	277.5	38.8	125	13	Q9PSM8	Q9psm8 echis carin
25	277	38.7	146	13	Q9DEF8	Q9def8 agkistrodon
26	274	38.3	146	13	Q8JIW1	Q8jiw1 agkistrodon
27	273	38.2	146	13	Q9YI92	Q9yi92 agkistrodon
28	272.5	38.1	145	13	Q9YGN4	Q9ygn4 agkistrodon
29	271.5	38.0	149	13	Q8UVC7	Q8uvc7 agkistrodon
30	268.5	37.6	149	13	Q8AYA3	Q8aya3 agkistrodon
31	256	35.8	123	13	Q9PSM5	Q9psm5 bothrops ja
32	256	35.8	155	13	Q8JIV8	Q8jiv8 agkistrodon
33	247	34.5	146	13	Q9DEA1	Q9dea1 agkistrodon
34	246	34.4	146	13	Q8JGT6	Q8jgt6 trimeresuru
35	239	33.4	124	13	Q90WL9	Q90wl9 agkistrodon
36	232	32.4	148	13	Q8AV98	Q8av98 trimeresuru
37	227.5	31.8	151	13	Q8JIW2	Q8jiw2 agkistrodon
38	215.5	30.1	158	13	Q90WI7	Q90wi7 bungarus fa
39	214.5	30.0	135	13	Q9PSM4	Q9psm4 lachesis mu
40	206.5	28.9	158	13	Q90WI6	Q90wi6 bungarus mu
41	205.5	28.7	158	13	Q90WI8	Q90wi8 bungarus fa
42	204.5	28.6	135	13	Q9PSN0	Q9psn0 bitis ariet
43	201.5	28.2	1479	4	Q9Y5P9	Q9y5p9 homo sapien
44	201.5	28.2	1479	4	Q9UBG0	Q9ubg0 homo sapien
45	199.5	27.9	719	6	O62623	O62623 bos taurus

ALIGNMENTS

RESULT 1
Q9IAM1
ID Q9IAM1 PRELIMINARY; PRT; 152 AA.
AC Q9IAM1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Agkisacutacin A chain.
OS Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Deinagkistrodon.
OX NCBI_TaxID=36307;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom;
RA Yu H.-X., Xiang K.-J., Liu J.;
RT "cDNA sequencing and analysis of eleven C-type lectin-like protein
RT subunits from Agkistrodon acutus.";
RL Chin. J. Biochem. Biophys. 0:0-0(2002).
DR EMBL; AF176420; AAF26286.2; -.
DR HSSP; P23806; 1IXX.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 152 AA; 17109 MW; 76A0F636DBF0D7AB CRC64;

Query Match 98.0%; Score 701; DB 13; Length 152;
Best Local Similarity 98.4%; Pred. No. 2.6e-66;
Matches 127; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 DCSSGWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
```

```
Db         84 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISYENWIEEESKKCLGVHIETGFHKWENFYCE 143

Qy        121 QQDPFVCEA 129
              |||||||||
Db        144 QQDPFVCEA 152


RESULT 2
Q9DEF9
ID   Q9DEF9      PRELIMINARY;      PRT;   152 AA.
AC   Q9DEF9;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Anticoagulant protein A precursor.
GN   ACP-A.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Ogawa T., Tani A.;
RT   "Purification, Sequencing, and cDNA Cloning of a Heterodimeric
RT   Anticoagulant Protein from Agkistrodon actus Venom Gland.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB036880; BAA99281.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
SQ   SEQUENCE   152 AA;  17124 MW;  5C59C0F503A4D223 CRC64;

  Query Match             91.9%;  Score 657;  DB 13;  Length 152;
  Best Local Similarity   92.2%;  Pred. No. 1.2e-61;
  Matches  119;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              |||| |||||||||| ||||||| ||||||||||||||||||||||||||  |||||||
Db         24 DCSSSWSSYEGHCYKAFKQSKTWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              |||||||||||||||||||||||||||||| ||||||||||||||  ||| ||||||||
Db         84 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCE 143

Qy        121 QQDPFVCEA 129
              |:|||||||
Db        144 QRDPFVCEA 152


RESULT 3
Q8JIW0
ID   Q8JIW0      PRELIMINARY;      PRT;   152 AA.
AC   Q8JIW0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ACF 1/2 A-chain.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yu H., Xiang K., Wang Y., Liu J.;
RT   "A chain of ACF 1/2 from Deinagkistrodon acutus.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY091758; AAM22786.1; -.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   152 AA;  17108 MW;  B0870F3DA0A4D23B CRC64;

  Query Match             91.9%;  Score 657;  DB 13;  Length 152;
  Best Local Similarity   92.2%;  Pred. No. 1.2e-61;
  Matches  119;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              |||| |||||||||| ||||||| ||||||||||||||||||||||||||  |||||||
Db         24 DCSSSWSSYEGHCYKAFKQSKTWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              |||||||||||||||||||||||||||||| ||||||||||||||  ||| ||||||||
Db         84 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCE 143

Qy        121 QQDPFVCEA 129
              |:|||||||
Db        144 QRDPFVCEA 152


RESULT 4
Q98UJ0
ID   Q98UJ0      PRELIMINARY;      PRT;   146 AA.
AC   Q98UJ0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Factor XI/factor X binding protein A chain.
GN   XI/XBP-A.
OS   Trimeresurus flavoviridis (Habu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=88087;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Tani A., Ogawa T., Nose T., Nikandorov N.N., Deshimaru M., Chijiwa T.,
RA   Chang C., Fukumaki Y., Ohno M.;
RT   "Characterization and molecular evolution of an anticoagulant protein
RT   from Agkistrodon actus venom.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB046491; BAB21452.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   146 AA;  16461 MW;  6ADF0E032544316C CRC64;

  Query Match             69.4%;  Score 496;  DB 13;  Length 146;
  Best Local Similarity   67.4%;  Pred. No. 1.2e-44;
  Matches   87;  Conservative   11;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| |:: ||| |||  ||:|  | ||||||||||||| ||: | :|
Db         18 DCLSGWSSYEGHCYKAFEKYKTWEDAERVCTEQAKGAHLVSIESSGEADFVAQLVTQNMK 77

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                ::||||| | | |||: |||||||:| ||||| ||| |||:  || | ||| | ||
Db         78 RLDFYIWIGLRVQGKVKQCNSEWSDGSSVSYENWIEAESKTCLGLEKETDFRKWVNIYCG 137
```

```
Qy        121 QQDPFVCEA 129
              ||:||||||
Db        138 QQNPFVCEA 146


RESULT 5
Q9DG39
ID   Q9DG39      PRELIMINARY;      PRT;   152 AA.
AC   Q9DG39;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Halyxin A-chain precursor.
GN   HXNA.
OS   Agkistrodon halys pallas (Chinese water mocassin) (Gloydius halys
OS   pallas).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Gloydius.
OX   NCBI_TaxID=8714;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koo B.H., Sohn Y.D., Kim D.S., Jang Y.S., Chung K.H.;
RT   "A novel coagulation factor Xa inhibitor from Korean snake
RT   (Agkistrodon halys) venom.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF190827; AAG17178.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    152       HALYXIN A-CHAIN.
SQ   SEQUENCE   152 AA;  17455 MW;  BDD74D1DC280C28D CRC64;

  Query Match             66.4%;  Score 475;  DB 13;  Length 152;
  Best Local Similarity   63.6%;  Pred. No. 2.1e-42;
  Matches   82;  Conservative   17;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              || | |||||||||| :|   ||| :|| || ||| | |||||||| ||||| ||::: :|
Db         24 DCPSGWSSYEGHCYNIFHLFKTWAEAERFCRKQVKGAHLVSIESSEEADFVAQLVSENMK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
                |::||||| : |:|||| :|||||:| :|||| ||| |||:  || | || | ||
Db         84 RYGIYIWIGLRVRGKKKQCSSQWSDGSSVSYQNWIEAESKTCLGLQKETEFRKWFNIYCG 143

Qy        121 QQDPFVCEA 129
              :::||||||
Db        144 ERNPFVCEA 152


RESULT 6
Q9YGN5
ID   Q9YGN5      PRELIMINARY;      PRT;   154 AA.
AC   Q9YGN5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Fibrinogen clotting inhibitor A chain.
OS   Agkistrodon halys brevicaudus (Korean slamosa snake) (Gloydius halys
OS   brevicaudus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Gloydius.
OX   NCBI_TaxID=66175;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim D.S., Koh Y.S.;
RT   "Purification and molecular cloning of snake venom fibrin clotting
RT   inhibitor.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125309; AAD18055.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   154 AA;  17293 MW;  8D06F7DDFA0D140D CRC64;

  Query Match             64.3%;  Score 460;  DB 13;  Length 154;
  Best Local Similarity   60.2%;  Pred. No. 8.1e-41;
  Matches   77;  Conservative   20;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKS 61
              | | | |  ||||: | |  || ||| ||: |  ||||||||:  ||||| ::|::|::
Db         27 CPSGWGSNNGHCYQAFNQRMTWEDAERFCSAQAKGGHLVSIETRAEADFVAHVVAERIET 86

Qy         62 AKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCEQ 121
              :  ||||||| : ||:||| |||||||:| ||||| ||| |||: ::: :||| | || |
Db         87 SFPHVWIGLRDEGKEQQCSSEWSDGSSVSYENWIEAESKTCLGLELDSNYHKWVNVYCGQ 146

Qy        122 QDPFVCEA 129
              ::||||||
Db        147 RNPFVCEA 154


RESULT 7
Q8UVC6
ID   Q8UVC6      PRELIMINARY;      PRT;   158 AA.
AC   Q8UVC6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Antithrombin 1 A chain.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yu H., Xiang K., Liu J.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF463522; AAL66391.1; -.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   158 AA;  17950 MW;  A73A9CB95997BFD7 CRC64;

  Query Match             61.4%;  Score 439;  DB 13;  Length 158;
  Best Local Similarity   60.2%;  Pred. No. 1.4e-38;
  Matches   77;  Conservative   18;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|   ||:|: :||:| |: | | ||| |||:| :||||||||| || ||| ||::| |:
Db         26 NCPPGWSAYDQYCYQVIKEPKNWDDAERFCTEQADGGHLVSIESKGERDFVAQLVSQSIE 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | : ||| ||| |||||||| |||||||:| || :|   :||  :  ||||||| |  |
Db         86 SVEDHVWTGLRVQNKEKQCSTEWSDGSSVSYENLLELYMRKCGALERETGFHKWINLGCI 145

Qy        121 QQDPFVCE 128
              | :||||:
Db        146 QLNPFVCK 153
```

```
RESULT 8
Q8AYA5
ID   Q8AYA5      PRELIMINARY;      PRT;   158 AA.
AC   Q8AYA5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Agglucetin-alpha 2 subunit precursor.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21542539; PubMed=11686327;
RA   Wang W.J., Huang T.F.;
RT   "A novel tetrameric venom protein, agglucetin from Agkistrodon acutus,
RT   acts as a glycoprotein Ib agonist.";
RL   Thromb. Haemost. 86:1077-1086(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Wang W.-J., Ling Q.-D., Huang T.-F.;
RT   "Molecular structure and functional characterization of agglucetin, a
RT   tetrameric glycoprotein Ib-binding protein, from Formosan pit viper.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF540646; AAN23125.1; -.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    158       AGGLUCETIN-ALPHA 2 SUBUNIT.
SQ   SEQUENCE   158 AA;  17977 MW;  9D0A9C936ABEA5E4 CRC64;

  Query Match             61.4%;  Score 439;  DB 13;  Length 158;
  Best Local Similarity   60.2%;  Pred. No. 1.4e-38;
  Matches   77;  Conservative   18;  Mismatches   33;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              :|    ||:|: :||:| |: | | ||| |||:| :||||||||| || ||| ||::| |:
Db         26 NCPPGWSAYDQYCYQVIKEPKNWDDAERFCTEQADGGHLVSIESKGERDFVAQLVSQNIE 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              | : ||| ||| |||||||| |||||||:| || :|   :||  :  |||||||  |  |
Db         86 SVEDHVWTGLRVQNKEKQCSTEWSDGSSVSYENLLELYMRKCGALERETGFHKWINLGCI 145

Qy        121 QQDPFVCE 128
              | :||||:
Db        146 QLNPFVCK 153


RESULT 9
Q9I841
ID   Q9I841      PRELIMINARY;      PRT;   136 AA.
AC   Q9I841;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Aggretin alpha chain (Fragment).
OS   Agkistrodon rhodostoma (Malayan pit viper) (Calloselasma rhodostoma).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Calloselasma.
OX   NCBI_TaxID=8717;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=99443731; PubMed=10512747;
RA   Chung C.H., Au L.C., Huang T.F.;
RT   "Molecular cloning and sequence analysis of aggretin, a collagen-like
RT   platelet aggregation inducer.";
RL   Biochem. Biophys. Res. Commun. 263:723-727(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Chung C.H., Au L.C., Huang T.F.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF244900; AAF79952.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER       1      1
FT   NON_TER     136    136
SQ   SEQUENCE   136 AA;  15796 MW;  3B474A4149F0027A CRC64;

  Query Match             58.8%;  Score 420.5;  DB 13;  Length 136;
  Best Local Similarity   59.7%;  Pred. No. 1e-36;
  Matches   77;  Conservative   16;  Mismatches   35;  Indels    1;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||    || |: |||: | : ||| :|| ||   | || || ||||:||||||  ||:|| :
Db          4 DCDFGWSPYDQHCYQAFNEQKTWDEAEKFCRAQENGAHLASIESNGEADFVSWLISQKDE 63

Qy         61 SA-KIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYC 119
              | : :||||||||||:||| |||||||:| || |:  :|||  :    ||| || |:||
Db         64 LADEDYVWIGLRAQNKEQQCSSEWSDGSSVSYENLIDLHTKKCGALEKLTGFRKWVNYYC 123

Qy        120 EQQDPFVCE 128
              ||   |||:
Db        124 EQMHAFVCK 132


RESULT 10
Q8JGT7
ID   Q8JGT7      PRELIMINARY;      PRT;   158 AA.
AC   Q8JGT7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   C-type lectin-like protein TMVA A chain.
OS   Trimeresurus mucrosquamatus (Taiwan habu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Protobothrops.
OX   NCBI_TaxID=103944;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wei Q., Lu Q.M., Jin Y., Li R., Wei J., Wang W., Xiong Y.;
RT   "Purification and cloning of a novel C-type lectin-like protein with
RT   platelet aggregation activity from Trimeresurus mucrosquamatus
RT   venom.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY099321; AAM43808.1; -.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Lectin.
SQ   SEQUENCE   158 AA;  18101 MW;  CD0B71F02289AA6E CRC64;

  Query Match             55.1%;  Score 394;  DB 13;  Length 158;
  Best Local Similarity   53.9%;  Pred. No. 8e-34;
  Matches   69;  Conservative   21;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||    ||:|: :||: | : | | |||||| : |   ||||||||||| ||| ||:|:|||
Db         26 DCIPGWSAYDRYCYQAFSEPKNWEDAESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIK 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              ::  :|||||| ||||:||  |||| ||:: ||  :: ||||  :   |    | | ||
Db         86 TSFQYVWIGLRIQNKEQQCRSEWSDASSVNYENLFKQSSKKCYALKKGTELRTWFNVYCG 145
```

```
Qy        121 QQDPFVCE 128
              :::||||:
Db        146 RENPFVCK 153


RESULT 11
Q8AV97
ID   Q8AV97      PRELIMINARY;      PRT;   158 AA.
AC   Q8AV97;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Flavocetin-A alpha chain.
OS   Trimeresurus flavoviridis (Habu).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Trimeresurus.
OX   NCBI_TaxID=88087;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=20402624; PubMed=10942790;
RA   Shin Y., Okuyama I., Hasegawa J., Morita T.;
RT   "Molecular cloning of glycoprotein Ib-binding protein, flavocetin-A,
RT   which inhibits platelet aggregation.";
RL   Thromb. Res. 99:239-247(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RA   Shin Y., Okuyama I., Hasegawa J., Morita T.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY149341; AAN72438.1; -.
SQ   SEQUENCE   158 AA;  18073 MW;  8C138650665CA454 CRC64;

  Query Match             54.7%;  Score 391;  DB 13;  Length 158;
  Best Local Similarity   53.9%;  Pred. No. 1.7e-33;
  Matches   69;  Conservative   21;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||   ||:|: :||: | : | | |||||| : |   ||||||||||| ||| ||:|:|||
Db         26 DCIPGWSAYDRYCYQAFSKPKNWEDAESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIK 85

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
              ::  :|||||| ||||:||  |||| ||:: || ::: |||| :   |    | | ||
Db         86 TSFQYVWIGLRIQNKEQQCRSEWSDASSVNYENLVKQFSKKCYALKKGTELRTWFNVYCG 145

Qy        121 QQDPFVCE 128
              ::||||:
Db        146 TENPFVCK 153


RESULT 12
Q9PSM9
ID   Q9PSM9      PRELIMINARY;      PRT;   131 AA.
AC   Q9PSM9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ECLV IX/X-BP alpha SUBUNIT=COAGULATION factor IX/factor X-binding
DE   protein alpha subunit.
OS   Echis carinatus (Saw-scaled viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Echis.
OX   NCBI_TaxID=40353;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96196635; PubMed=8611513;
RA   Chen Y.L., Tsai I.H.;
RT   "Functional and sequence characterization of coagulation factor
RT   IX/factor X-binding protein from the venom of Echis carinatus
RT   leucogaster.";
RL   Biochemistry 35:5264-5271(1996).
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   131 AA;  15439 MW;  B85E6C5CBF317E24 CRC64;

  Query Match             54.5%;  Score 390;  DB 13;  Length 131;
  Best Local Similarity   54.6%;  Pred. No. 1.7e-33;
  Matches   71;  Conservative   17;  Mismatches   40;  Indels    2;  Gaps    2;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKI- 59
              ||   |||:|||||||| : ||| ||| || ||   |||||:||| | ||| :||:: :
Db          1 DCLPGWSSHEGHCYKVFNEYKTWKDAEKFCKKQGKSGHLVSVESSEEGDFVAKLISENLE 60

Qy         60 KSAKI-HVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFY 118
              ||  |  || ||  : : |||| |||||| |  : | ::: :||||:  :| | || | |
Db         61 KSHSIDFVWTGLTYKGRWKQCSSEWSDGSKIKYQKWGKQQPRKCLGLEKQTEFRKWVNLY 120

Qy        119 CEQQDPFVCE 128
              ||:   | ||
Db        121 CEEPQRFTCE 130


RESULT 13
Q9PSM6
ID   Q9PSM6      PRELIMINARY;      PRT;   142 AA.
AC   Q9PSM6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Platelet glycoprotein IB-binding protein alpha subunit, GPIB-BP alpha
DE   subunit.
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96209990; PubMed=8631868;
RA   Kawasaki T., Fujimura Y., Usami Y., Suzuki M., Miura S., Sakurai Y.,
RA   Makita K., Taniuchi Y., Hirano K., Titani K.;
RT   "Complete amino acid sequence and identification of the platelet
RT   glycoprotein Ib-binding site of jararaca GPIb-BP, a snake venom
RT   protein isolated from Bothrops jararaca.";
RL   J. Biol. Chem. 271:10635-10639(1996).
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   142 AA;  16720 MW;  E28EFDEBEF922004 CRC64;

  Query Match             53.5%;  Score 382.5;  DB 13;  Length 142;
  Best Local Similarity   51.1%;  Pred. No. 1.2e-32;
  Matches   68;  Conservative   23;  Mismatches   37;  Indels    5;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWT-----DAESFCTKQVNGGHLVSIESSGEADFVGQLI 55
              :| ||||::  :||| |:| ::|      ||| ||::|  ||||||||  ||||| ||:
Db          5 ECPSDWSTHRQYCYKFFQQKESWDDRSEYDAERFCSEQAKGGHLVSIESDEEADFVAQLV 64

Qy         56 AQKIKSAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWE 115
              |  |  :| :|||||| :||::||| :||| ||:| || :    |||  :  : || ||
Db         65 APNIGKSKYYVWIGLRIENKKQQCSSKWSDYSSVSYENLVRGNVKKCFALEKKQGFRKWV 124

Qy        116 NFYCEQQDPFVCE 128
              |  | : :||||:
Db        125 NIDCVEGNPFVCK 137
```

```
RESULT 14
Q8JIV6
ID   Q8JIV6      PRELIMINARY;      PRT;   155 AA.
AC   Q8JIV6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Antithrombin A A-chain.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yu H., Xiang K., Wang Y., Liu J.;
RT   "A chain of antithrombin A from Deinagkistrodon acutus.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY091762; AAM22790.1; -.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ   SEQUENCE   155 AA;  17715 MW;  66E55C95E997A1D7 CRC64;

  Query Match             53.2%;  Score 380.5;  DB 13;  Length 155;
  Best Local Similarity   54.7%;  Pred. No. 2.1e-32;
  Matches   70;  Conservative   18;  Mismatches   39;  Indels    1;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||   || |   ||: ||  ||| ||| |||:| ||||||| ||:||||||  :::: ||
Db         24 DCLPGWSCYIRFCYQPFKLLKTWEDAERFCTEQANGGHLVSFESAGEADFVAGVLSENIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               | :||||||| ||: :||| :||| | :| || :|  |||| :  :|||  ||| ||
Db         84 -IKPYVWIGLRVQNEGQQCSSKWSDSSKVSYENLVEPFSKKCFVLKKDTGFRTWENVYCG 142

Qy        121 QQDPFVCE 128
                :  |:|:
Db        143 LKHVFMCK 150


RESULT 15
Q9DEA2
ID   Q9DEA2      PRELIMINARY;      PRT;   155 AA.
AC   Q9DEA2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Agkicetin alpha subunit precursor.
GN   GPIBA1.
OS   Agkistrodon acutus (Hundred-pace snake) (Deinagkistrodon acutus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Deinagkistrodon.
OX   NCBI_TaxID=36307;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chen Y.L., Hong T.M., Chang T., Tsai I.H.;
RT   "cDNA sequence and functional characterization of glycoprotein Ib-
RT   binding protein from the venom of Deinagkistrodon acutus.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF102901; AAG42040.1; -.
DR   HSSP; P23806; 1IXX.
DR   InterPro; IPR001304; Lectin_C.
DR   InterPro; IPR003990; Pancreatis_ac.
DR   Pfam; PF00059; lectin_c; 1.
DR   PRINTS; PR01504; PNCREATITSAP.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    155       AGKICETIN ALPHA SUBUNIT.
SQ   SEQUENCE   155 AA;  17798 MW;  92B40281A797AACD CRC64;

  Query Match             52.8%;  Score 377.5;  DB 13;  Length 155;
  Best Local Similarity   54.7%;  Pred. No. 4.3e-32;
  Matches   70;  Conservative   18;  Mismatches   39;  Indels    1;  Gaps    1;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIK 60
              ||   ||||   ||: ||  ||| ||| |||:| ||||||| ||: |||||  :::: ||
Db         24 DCLPGWSSYIRFCYQPFKLLKTWEDAERFCTEQANGGHLVSFESAREADFVAGVLSENIK 83

Qy         61 SAKIHVWIGLRAQNKEKQCSIEWSDGSSISKENWIEEESKKCLGVHIETGFHKWENFYCE 120
               | :|||||| ||: :||| :||| | :| || :|  |||| :  :|||  ||| ||
Db         84 -IKPYVWIGLRVQNEGQQCSSKWSDSSKVSYENLVEPFSKKCFVLKKDTGFRTWENVYCG 142

Qy        121 QQDPFVCE 128
                :  |:|:
Db        143 LKHVFMCK 150


Search completed: December  8, 2003, 09:54:48
Job time : 64.6571 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:34:19 ; Search time 3342.94 Seconds
(without alignments)
937.879 Million cell updates/sec

Title: US-09-938-114-2
Perfect score: 715
Sequence: 1 DCSSDWSSYEGHCYKVFKQS.....GFHKWENFYCEQQDPFVCEA 129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22859/app_query.fasta_1.725
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_4382_@runat_08122003_091000_22859 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	382.5	53.5	556	12	BM401668	BM401668 JL2H08F S
2	373.5	52.2	501	12	BM401648	BM401648 JL2B03F S
3	359	50.2	460	12	BM401460	BM401460 JH1D06F S
4	354.5	49.6	618	12	BM401682	BM401682 PH012F Sn
5	342.5	47.9	458	12	BM401641	BM401641 JL1H06F S
6	341.5	47.8	511	12	BM401631	BM401631 JL1F11F S
7	331.5	46.4	406	12	BM401598	BM401598 JH4F11F S
8	324.5	45.4	410	12	BM401605	BM401605 JH4H07F S
9	323	45.2	418	12	BM401662	BM401662 JL2G06F S
10	312.5	43.7	440	12	BM401459	BM401459 JH1D05F S
11	312.5	43.7	453	12	BM401653	BM401653 JL2D01F S
c 12	312.5	43.7	461	12	BM401683	BM401683 PH012R Sn
13	300.5	42.0	440	12	BM401637	BM401637 JL1G11F S
14	300.5	42.0	456	12	BM401659	BM401659 JL2F01F S
15	279	39.0	654	12	BM401401	BM401401 GH025F Sn
16	277	38.7	574	12	BM401417	BM401417 GH061F Sn
17	248.5	34.8	405	12	BM401607	BM401607 JH4H10F S
18	235	32.9	510	12	BM401667	BM401667 JL2H07F S
19	235	32.9	512	12	BM401495	BM401495 JH2B09F S
20	232.5	32.5	497	12	BM401552	BM401552 JH3E12F S
21	230	32.2	633	12	BM401686	BM401686 PH015FB S
22	226	31.6	504	12	BM401628	BM401628 JL1E12F S
23	223.5	31.3	388	12	BM401625	BM401625 JL1E05F S
24	223.5	31.3	475	12	BM401621	BM401621 JL1D02F S
25	223	31.2	378	12	BM401618	BM401618 JL1C07F S
26	216	30.2	378	12	BM401620	BM401620 JL1C11F S
27	216	30.2	391	12	BM401469	BM401469 JH1E09F S
28	211.5	29.6	365	12	BM401665	BM401665 JL2H01F S
29	210.5	29.4	453	14	CA843095	CA843095 ir53d01.y
30	210.5	29.4	535	14	CA772255	CA772255 io93g12.y
31	209	29.2	301	12	BM401654	BM401654 JL2D07F S
32	209	29.2	374	12	BM401520	BM401520 JH2G02F S
33	208.5	29.2	486	14	CA773380	CA773380 im63c12.y
34	208.5	29.2	546	13	BQ776439	BQ776439 il32f09.y
c 35	208.5	29.2	642	12	BM312458	BM312458 ig78g11.x
36	207.5	29.0	570	14	CA842080	CA842080 ip28a05.y
c 37	207.5	29.0	628	13	BU072359	BU072359 im45h03.x
38	206.5	28.9	425	12	BM352964	BM352964 ig69f02.y
39	206.5	28.9	551	14	CA775597	CA775597 io87g10.y
c 40	206.5	28.9	593	14	CA774835	CA774835 ip13b07.x
41	205.5	28.7	453	14	CA868012	CA868012 ir76d01.y
42	205.5	28.7	546	13	BQ778069	BQ778069 il40f01.y
43	205.5	28.7	547	14	CA772731	CA772731 io84b10.y
44	205.5	28.7	550	12	BM313297	BM313297 ig82d01.y
45	205.5	28.7	556	12	BM313209	BM313209 ig80h07.y

ALIGNMENTS

RESULT 1
BM401668
LOCUS BM401668 556 bp mRNA linear EST 01-MAY-2002
DEFINITION JL2H08F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401668
VERSION BM401668.1 GI:20376296
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.

```
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      146 a    126 c    154 g    129 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                4.82e-34        Length:          556
Score:                    382.50          Matches:         68
Percent Similarity:       68.42%          Conservative:    23
Best Local Similarity:    51.13%          Mismatches:      37
Query Match:              53.50%          Indels:          5
DB:                       12              Gaps:            1

US-09-938-114-2 (1-129) x BM401668 (1-556)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||||||||||:::::      :::|||||||||   |||:::|||
Db         82 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 141

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              :::::::|||                 |||||||||   ||||||:::::::|||
Db        142 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 201

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||||||||      ||||||||||||||   ||||||:::
Db        202 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 261

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::||||||||||||||||||   :::||||||
Db        262 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 321

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||   :::
Db        322 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 381

Qy         96 GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
                        |||||||||      :::      :::   ||||||   ||||||
Db        382 CGAGGAAATGTGAAAAAGTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCGTAAGTGGGTC 441

Qy        116 AsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
              |||      |||   :::   :::||||||||||||:::
Db        442 AATATTGACTGTGTAGAAGGAAATCCTTTCGTGTGCAAG 480

RESULT 2
BM401648
LOCUS       BM401648                501 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2B03F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401648
VERSION     BM401648.1  GI:20376276
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      134 a    102 c    139 g    122 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:                4.55e-33        Length:          501
Score:                    373.50          Matches:         67
Percent Similarity:       66.42%          Conservative:    22
Best Local Similarity:    50.00%          Mismatches:      40
Query Match:              52.24%          Indels:          5
DB:                       12              Gaps:            1

US-09-938-114-2 (1-129) x BM401648 (1-501)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||||||||||:::::      :::|||||||||   |||:::|||
Db         84 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 143

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                 |||||||||   ||||||:::::::|||
Db        144 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 203
```

```
Qy          36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
               ||||||||||||||||||||||||||      ||||||||||||||||   ||||||:::
Db         204 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 263

Qy          56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
               |||      |||      :::|||   :::||||||||||||||||||   :::||||||
Db         264 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 323

Qy          76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
               ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||   :::
Db         324 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 383

Qy          96 GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
                         |||||||||      :::      :::   ||||||   ||||||
Db         384 CGAGGAAATGTGAAAAAGTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCNTAAGTGGGTC 443

Qy         116 AsnPheTyrCysGluGlnGlnAspProPheValCysGluAla 129
               |||      |||         :::||||||   |||   :::
Db         444 AATATTGACTGTGTANAAGGAAATCCTTTCNTGTGCAATTCA 485
```

RESULT 3
BM401460
LOCUS BM401460 460 bp mRNA linear EST 01-MAY-2002
DEFINITION JH1D06F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401460
VERSION BM401460.1 GI:20376088
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 460)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL13A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .460
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 106 a 106 c 145 g 103 t
ORIGIN

```
Alignment Scores:
Pred. No.:              1.91e-31        Length:         460
Score:                  359.00          Matches:        61
Percent Similarity:     74.51%          Conservative:   15
Best Local Similarity:  59.80%          Mismatches:     26
Query Match:            50.21%          Indels:         0
DB:                     12              Gaps:           0

US-09-938-114-2 (1-129) x BM401460 (1-460)

Qy           1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
               ||||||      ||||||||||||||||||||   |||||:::||||||   :::
Db         143 GATTGTCCCCCTGATTGGTCCTCCTATGAAGGGAGTTGCTACAGGGTCTTCACTGAACCG 202

Qy          21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
               :::   |||   |||||||||   |||||:::::|||       ||||||||||||||
Db         203 CAAAACTGGGCCGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTC 262

Qy          41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
               ||||||||||||   ||||||||||||||||||   ||||||:::|||:::::::|||
Db         263 TCTATCGAAAGCGAAGGAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTGAGAGGATAGAC 322

Qy          61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
                  :::   ||||||||||||||||||||   :::::::|||   :::||||||||||
Db         323 AGACCCCAATCCCATGTCTGGATCGGACTGAGGGGTGAACACAAAGGACAGCAATGCAGC 382

Qy          81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
                  ||||||||||||||   :::      ||||||   :::|||      :::|||
Db         383 TCGAACTGGAGCGATGGCTCCTGCGTCTGTTATGAGAACCTGGTTGAACGAAGAACGAAA 442

Qy         101 LysCys 102
               ||||||
Db         443 AAATGT 448
```

RESULT 4
BM401682
LOCUS BM401682 618 bp mRNA linear EST 01-MAY-2002
DEFINITION PH012F Snake Bothrops insularis library IL2 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401682
VERSION BM401682.1 GI:20376310
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 618)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL09A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .618
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL2"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 157 a 124 c 180 g 149 t 8 others
ORIGIN

Alignment Scores:

Pred. No.:	9.71e-31	Length:	618
Score:	354.50	Matches:	72
Percent Similarity:	71.09%	Conservative:	19
Best Local Similarity:	56.25%	Mismatches:	36
Query Match:	49.58%	Indels:	4
DB:	12	Gaps:	1

US-09-938-114-2 (1-129) x BM401682 (1-618)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||:::      ||||||||||||||||||      |||:::   ||||||::::::
Db        214 GATTGTAACCCTTATTGGTCCTCCTATGAAGGGCGTTCCTACAGGTTCTTCAAAGAAGCG 273

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||   |||   ||||||       ||||||:::::::::|||||||||||||||||||
Db        274 AAAAACTGGGACGATGCANAGAGGTTCTGCTCGGAGGAGGTGAATGGCGGCCATCTGGTC 333

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db        334 TCTATCGAAAGCGCCGGAGAAGCAGCCTTTGTGGCCCAGCTGGTCGCTGAGAACAAACAC 393

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||   :::::::::|||||||||||||:::|||   |||   |||   |||||||||
Db        394 AGTCGCGAATTGAATGTCTGGATCGGAATGAGGGTTCAAGGCAAAGA-AAACAATGCANC 452

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLys 100
              ||||||||||||||||   :::|||   ||||||||||||:::          |||
Db        453 ---TCGTGGAGCGATGGCTCCAC-GTCAGTTATGAGAACTGGATTAAAGCANAATTAAAA 508

Qy        101 LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGlu 120
                ||:::|||:::     :::   ||||||   ||||||   |||      ||||||
Db        509 ACATGTATTGGGTTGAAAAAAGATGCAGGGTTTCNTAAGTGGAGCAATCTTAACTGTGAA 568

Qy        121 GlnGlnAspProPheValCysGlu 128
              |||      ||||||  ||||||
Db        569 CAACNAATTCCTTTCTC-TGCGAG 591
```

RESULT 5
BM401641
LOCUS BM401641 458 bp mRNA linear EST 01-MAY-2002
DEFINITION JL1H06F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401641
VERSION BM401641.1 GI:20376269
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 458)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .458
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 122 a 90 c 130 g 111 t 5 others
ORIGIN

Alignment Scores:

Pred. No.:	1.53e-29	Length:	458
Score:	342.50	Matches:	62
Percent Similarity:	66.13%	Conservative:	20
Best Local Similarity:	50.00%	Mismatches:	37
Query Match:	47.90%	Indels:	5
DB:	12	Gaps:	1

US-09-938-114-2 (1-129) x BM401641 (1-458)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   ||||||||||||:::::      :::|||||||||   |||:::|||
Db         84 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 143

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                 ||||||       ||||||::::::|||
Db        144 GAGAGCTGGGATGATGTGAGCTGGTATGATGCANANAGGTTCTGCTCGGAGCAGGCGAAG 203

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||||||||      ||||||||||||||   |||||||:::
Db        204 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 263

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::|||||||||||||||||   :::||||||
Db        264 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 323

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||   :::
Db        324 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 383

Qy         96 GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
                        |||||||||      :::      :::      |||   ||||||
Db        384 CGAGGAAATGTGAAAAAGTGTTTTGCGCTGGAAAAAAAGCAANGGTTTCNTAAGTGGGTC 443

Qy        116 AsnPheTyrCys 119
              |||      |||
Db        444 AATATTGACTGT 455
```

RESULT 6
BM401631
LOCUS BM401631 511 bp mRNA linear EST 01-MAY-2002
DEFINITION JL1F11F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.

```
ACCESSION   BM401631
VERSION     BM401631.1  GI:20376259
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      135 a    105 c    149 g    119 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:                2.35e-29        Length:          511
Score:                    341.50          Matches:         62
Percent Similarity:       66.94%          Conservative:    19
Best Local Similarity:    51.24%          Mismatches:      35
Query Match:              47.76%          Indels:          5
DB:                       12              Gaps:            1

US-09-938-114-2 (1-129) x BM401631 (1-511)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||||||||||::::::      :::||||||||||   |||:::|||
Db        148 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 207

Qy         21 LysThrTrpThr---------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                 |||||||||   ||||||::::::|||
Db        208 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 267

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||||||      ||||||||||||||      |||:::
Db        268 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCANCTGGTC 327

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::||||||||||||||||   :::||||||
Db        328 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 387

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
              :::   |||||||||   :::||||||||||   ||||||:::|||   ||||||   :::
Db        388 AAACANCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 447

Qy         96 GluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGlu 115
                         |||||||||      :::      :::   ||||||   ||||||
Db        448 CGAGGAAATGTGAAAAAGTGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCNTAAGTGGGTC 507

Qy        116 Asn 116
              |||
Db        508 AAT 510

RESULT 7
BM401598
LOCUS       BM401598                 406 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JH4F11F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401598
VERSION     BM401598.1  GI:20376226
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL15A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT       95 a     91 c    126 g     94 t
ORIGIN

Alignment Scores:
Pred. No.:                2.41e-28        Length:          406
Score:                    331.50          Matches:         59
Percent Similarity:       78.49%          Conservative:    14
Best Local Similarity:    63.44%          Mismatches:      19
Query Match:              46.36%          Indels:          1
DB:                       12              Gaps:            1

US-09-938-114-2 (1-129) x BM401598 (1-406)
```

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||:::     ||||||||||||||||||     |||:::   ||||||
Db        129 GATTGTAACCCTTATTGGTCCTCCTATGAAGGGCGTTCCTACAGGTTCTTCAAAACACCG 188

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
              |||   |||    |||||||||   ||||||:::::::::|||||||||||||||||||||
Db        189 AAAAACTGGGACGATGCAGAGAGGTTCTGCTCGGAGGAGGTGAATGGCGGCCATCTGGTC 248

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              ||||||||||||:::|||||||||||||||||   ||||||:::   :::   |||:::
Db        249 TCTATCGAAAGCGCCGGAGAAGCAGACTTTGTGGCCCAGCTGGTCACTGAGAACATACAG 308

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              |||   ::::::   ||||||||||||::::::   |||   ||||||||||||||||||
Db        309 AGCCGCGAATTGTTTGTCTGGATCGGAATGAAGGTTCAAGGCAAAGAAAAGCAATGCAGC 368

Qy         81 IleGluTrpSerAspGlySerSerIleSerLysGluAsn 93
                    ||||||||||||||||||:::|||   ||||||
Db        369 ---TCGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAAC 404
```

RESULT 8
BM401605
LOCUS BM401605 410 bp mRNA linear EST 01-MAY-2002
DEFINITION JH4H07F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401605
VERSION BM401605.1 GI:20376233
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 410)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .410
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 106 a 87 c 120 g 97 t
ORIGIN

Alignment Scores:

Pred. No.:	1.58e-27	Length:	410
Score:	324.50	Matches:	59
Percent Similarity:	71.30%	Conservative:	18
Best Local Similarity:	54.63%	Mismatches:	26
Query Match:	45.38%	Indels:	5
DB:	12	Gaps:	1

US-09-938-114-2 (1-129) x BM401605 (1-410)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   ||||||||||||::::::     :::|||||||||   |||:::|||
Db         84 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 143

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                 |||||||||   ||||||::::::|||
Db        144 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 203

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||||||||       ||||||||||||||   ||||||:::
Db        204 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 263

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||   |||    :::|||   :::|||||||||||||||||   :::||||||
Db        264 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 323

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrpIle 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||   :::
Db        324 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 383

Qy         96 GluGluGluSerLysLysCysLeu 103
                        ||||||||||||
Db        384 CGAGGAAATGTGAAAAAGTGTTTG 407
```

RESULT 9
BM401662
LOCUS BM401662 418 bp mRNA linear EST 01-MAY-2002
DEFINITION JL2G06F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401662
VERSION BM401662.1 GI:20376290
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 418)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .418
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis

venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."

BASE COUNT 116 a 90 c 122 g 90 t
ORIGIN

Alignment Scores:

Pred. No.:	2.42e-27	Length:	418
Score:	323.00	Matches:	58
Percent Similarity:	70.09%	Conservative:	17
Best Local Similarity:	54.21%	Mismatches:	32
Query Match:	45.17%	Indels:	0
DB:	12	Gaps:	0

US-09-938-114-2 (1-129) x BM401662 (1-418)

```
Qy         22 ThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuValSer 41
              :::|||   |||||||||   ||||||::::::|||      ||||||||||||||||
Db         14 AGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAGGGCGGCCATCTGGTCTCT 73

Qy         42 IleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLysSer 61
              |||||||||      ||||||||||||||   ||||||:::|||      |||
Db         74 ATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTCGCTCCGAACATAGGGAAA 133

Qy         62 AlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSerIle 81
              :::|||   :::|||||||||||||||||||   :::|||||::::::|||||||||
Db        134 TCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAAAAACAGCAATGCAGCTCG 193

Qy         82 GluTrpSerAspGlySerSerIleSerLysGluAsnTrpIleGluGluGluSerLysLys 101
              :::|||||||||   ||||||:::|||   ||||||   :::             ||||||
Db        194 AAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTTCGAGGAAATGTGAAAAAG 253

Qy        102 CysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyrCysGluGln 121
              |||      :::      :::   ||||||   ||||||   |||      |||   :::
Db        254 TGTTTTGCGCTGGAAAAAAAGCAAGGGTTTCGTAAGTGGGTCAATATTGACTGTGTAGAA 313

Qy        122 GlnAspProPheValCysGlu 128
                 :::||||||||||||:::
Db        314 GGAAATCCTTTCGTGTGCAAG 334
```

RESULT 10
BM401459
LOCUS BM401459 440 bp mRNA linear EST 01-MAY-2002
DEFINITION JH1D05F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401459
VERSION BM401459.1 GI:20376087
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.
REFERENCE 1 (bases 1 to 440)
AUTHORS Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
TITLE A survey of gene expression and diversity in the venom glands of the pitviper snake Bothrops insularis through the generation of expressed sequence tags (ESTs)
JOURNAL Gene 299 (1-2), 279-291 (2002)
MEDLINE 22347338
PUBMED 12459276
COMMENT Contact: Paulo Lee Ho
Centro de Biotecnologia
Instituto Butantan
Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
Tel: 55 11 37 26 7222 ext. 2083
Fax: 55 11 37 26 1505
Email: hoplee@usp.br
This EST corresponds to cluster BITL07A (see Reference)
Seq primer: M13F.
FEATURES Location/Qualifiers
source 1. .440
/organism="Bothrops insularis"
/mol_type="mRNA"
/db_xref="taxon:8723"
/tissue_type="venom glands"
/clone_lib="Snake Bothrops insularis library IL3"
/note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis venom glands were primed with oligo-(dT) and reverse transcribed to cDNA using Superscript Plasmid System for cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs were selected by size (350-600 pb and up 600 pb) in agarose gel electrophoresis, linked to Eco RI adapters and directionally cloned in pGEM11Zf+ vector (Promega). ESTs were generated from random clones and grouped in unique sequences. The putative identification of each EST or cluster was obtained through Blast searches (e-value < e-05)."
BASE COUNT 113 a 96 c 130 g 101 t
ORIGIN

Alignment Scores:

Pred. No.:	4.26e-26	Length:	440
Score:	312.50	Matches:	58
Percent Similarity:	72.82%	Conservative:	17
Best Local Similarity:	56.31%	Mismatches:	22
Query Match:	43.71%	Indels:	6
DB:	12	Gaps:	1

US-09-938-114-2 (1-129) x BM401459 (1-440)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   ||||||||||||:::::::      :::|||||||||   |||:::|||
Db        134 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 193

Qy         21 LysThrTrpThr----------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                |||||||||   ||||||::::::|||
Db        194 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 253

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              |||||||||||||||||||||||||||      ||||||||||||||   ||||||:::
Db        254 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 313

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::||||||||||||||||||   :::||||||
Db        314 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 373

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn-TrpIl 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   |||||| |||
Db        374 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 433

Qy         95 eGluGlu 97
                ||||||
Db        434 CGAGGAA 440
```

RESULT 11
BM401653
LOCUS BM401653 453 bp mRNA linear EST 01-MAY-2002
DEFINITION JL2D01F Snake Bothrops insularis library IL3 Bothrops insularis cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION BM401653
VERSION BM401653.1 GI:20376281
KEYWORDS EST.
SOURCE Bothrops insularis (island jararaca)
ORGANISM Bothrops insularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea; Viperidae; Crotalinae; Bothrops.

```
REFERENCE   1  (bases 1 to 453)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .453
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      120 a     96 c    132 g    104 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                4.44e-26        Length:         453
Score:                    312.50          Matches:        58
Percent Similarity:       72.82%          Conservative:   17
Best Local Similarity:    56.31%          Mismatches:     22
Query Match:              43.71%          Indels:         6
DB:                       12              Gaps:           1

US-09-938-114-2 (1-129) x BM401653 (1-453)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   ||||||||||||::::::      :::|||||||||   |||:::|||
Db        135 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 194

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                |||||||||   ||||||::::::|||
Db        195 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 254

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||||||||      ||||||||||||||   ||||||:::
Db        255 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 314

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::||||||||||||||||||   :::||||||
Db        315 GCTCCNAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 374

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn-TrpIl 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||  |||
Db        375 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 434

Qy         95 eGluGlu 97
              |||||||
Db        435 CGAGGAA 441
```

```
RESULT 12
BM401683/c
LOCUS       BM401683                461 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  PH012R Snake Bothrops insularis library IL2 Bothrops insularis cDNA
            3' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401683
VERSION     BM401683.1  GI:20376311
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL09A (see Reference)
            Seq primer: M13R.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL2"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      104 a    125 c    107 g    120 t      5 others
ORIGIN

Alignment Scores:
Pred. No.:                4.56e-26        Length:         461
Score:                    312.50          Matches:        57
Percent Similarity:       74.47%          Conservative:   13
Best Local Similarity:    60.64%          Mismatches:     23
Query Match:              43.71%          Indels:         1
DB:                       12              Gaps:           1

US-09-938-114-2 (1-129) x BM401683 (1-461)

Qy         35 AsnGlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeu 54
              ||||||||||||||||||||||||||:::|||||||||   ||||||   ||||||
Db        459 AATGGCGGCCATCTGGTNTCTATCGAAAGCGCCGGAGAAGCAGCCTTTGTGGCCCAGCTG 400

Qy         55 IleAlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsn 74
              :::   :::         |||   :::::::::||||||||||:::|||   |||
Db        399 GTCGNTGAGAACAAACACAGTCGCGAATTGAATGTCTGGATCGGAATGAGGGTTCAAGGC 340

Qy         75 LysGluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsnTrp 94
              ||||||||||||||||||      ||||||||||||||||||:::|||   |||||||||
Db        339 AAAGAAAAGCAATGCAGC---TCGTGGAGCGATGGCTCCAGCGTCAGTTATGAGAACTGG 283
```

```
Qy         95 IleGluGluGluSerLysLysCysLeuGlyValHisIleGluThrGlyPheHisLysTrp 114
              |||:::   |||   |||   |||:::|||:::      :::   ||||||   ||||||
Db        282 ATTAAAGCAGAATTAAAAACATGTATTGGGTTGAAAAAAGATGCAGGGTTTCGTAAGTGG 223

Qy        115 GluAsnPheTyrCysGluGlnGlnAspProPheValCysGlu 128
                 |||      |||||||||:::   |||   |||||||||
Db        222 AGCAATNTTAACTGTGAACAACGAATTCCTTTNGTCTGCGAG 181

RESULT 13
BM401637
LOCUS       BM401637                 440 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL1G11F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401637
VERSION     BM401637.1  GI:20376265
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 440)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .440
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      110 a     99 c    129 g    100 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.04e-24        Length:         440
Score:                  300.50          Matches:        54
Percent Similarity:     72.45%          Conservative:   17
Best Local Similarity:  55.10%          Mismatches:     22
Query Match:            42.03%          Indels:         5
DB:                     12              Gaps:           1

US-09-938-114-2 (1-129) x BM401637 (1-440)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   ||||||||||||:::::      :::|||||||||   |||:::|||

Db        145 GAGTGTCCCTCTGATTGGTCCACCCATAGACAGTATTGCTACAAGTTCTTCCAACAAAAG 204

Qy         21 LysThrTrpThr--------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                ||||||||   ||||||:::::::|||

Db        205 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 264

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              |||||||||||||||||||||||||      ||||||||||||||   ||||||:::

Db        265 GGCGGCCATCTGGTCTCTATCGAAAGCGATGAAGAAGCAGACTTTGTGGCCCAGCTGGTC 324

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||      :::|||   :::|||||||||||||||||   :::||||||

Db        325 GCTCCGAACATAGGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 384

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn 93
              ::::::||||||||||   :::|||||||||   |||   :::|||   ||||||

Db        385 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCNGCGTCAGTTATGAGAAC 438

RESULT 14
BM401659
LOCUS       BM401659                 456 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  JL2F01F Snake Bothrops insularis library IL3 Bothrops insularis
            cDNA 5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401659
VERSION     BM401659.1  GI:20376287
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL07A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL3"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      117 a     100 c     131 g     104 t      4 others
ORIGIN

Alignment Scores:
Pred. No.:              1.09e-24        Length:         456
Score:                  300.50          Matches:        56
```

```
Percent Similarity:     70.87%          Conservative:  17
Best Local Similarity:  54.37%          Mismatches:    24
Query Match:            42.03%          Indels:        6
DB:                     12              Gaps:          1

US-09-938-114-2 (1-129) x BM401659 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||||||||||:::::::      :::|||||||||   |||:::|||
Db        149 GAGTGTCCCTCTGATTGGTCCACCCATANACAGTATTGCTACAAGTTCTTCCAACAAAAG 208

Qy         21 LysThrTrpThr---------------AspAlaGluSerPheCysThrLysGlnValAsn 35
              ::::::|||                 |||||||||   |||||:::::|||
Db        209 GAGAGCTGGGATGATGTGAGCTGGTATGATGCAGAGAGGTTCTGCTCGGAGCAGGCGAAG 268

Qy         36 GlyGlyHisLeuValSerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIle 55
              ||||||||||||||||||||   |||   |||||||||||||||      |||:::
Db        269 GGCGGCCATCTGGTCTCTATCNAAAGCGATGAAGAAGCAGACTTTGTGGCCCANCTGGTC 328

Qy         56 AlaGlnLysIleLysSerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLys 75
              |||      |||     :::|||   :::||||||||||||||||||   :::||||||
Db        329 GCTCCGAACATANGGAAATCCAAATACTATGTCTGGATCGGACTGAGGATTGAAAACAAA 388

Qy         76 GluLysGlnCysSerIleGluTrpSerAspGlySerSerIleSerLysGluAsn-TrpIl 95
              ::::::|||||||||   :::|||||||||   ||||||:::|||   ||||||  |||
Db        389 AAACAGCAATGCAGCTCGAAGTGGAGCGATTACTCCAGCGTCAGTTATGAGAACCTGGTT 448

Qy         95 eGluGlu 97
                ||||||
Db        449 CGAGGAA 455

RESULT 15
BM401401
LOCUS       BM401401                654 bp    mRNA    linear   EST 01-MAY-2002
DEFINITION  GH025F Snake Bothrops insularis library IL2 Bothrops insularis cDNA
            5' similar to Snake venom C-type lectin, mRNA sequence.
ACCESSION   BM401401
VERSION     BM401401.1  GI:20376029
KEYWORDS    EST.
SOURCE      Bothrops insularis (island jararaca)
  ORGANISM  Bothrops insularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Bothrops.
REFERENCE   1  (bases 1 to 654)
  AUTHORS   Junqueira-de-Azevedo,I.L.M. and Ho,P.L.
  TITLE     A survey of gene expression and diversity in the venom glands of
            the pitviper snake Bothrops insularis through the generation of
            expressed sequence tags (ESTs)
  JOURNAL   Gene 299 (1-2), 279-291 (2002)
  MEDLINE   22347338
   PUBMED   12459276
COMMENT     Contact: Paulo Lee Ho
            Centro de Biotecnologia
            Instituto Butantan
            Av. Vital Brazil, 1500, Sao Paulo SP, BRAZIL, 05503-900
            Tel: 55 11 37 26 7222 ext. 2083
            Fax: 55 11 37 26 1505
            Email: hoplee@usp.br
            This EST corresponds to cluster BITL10A (see Reference)
            Seq primer: M13F.
FEATURES             Location/Qualifiers
     source          1. .654
                     /organism="Bothrops insularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8723"
                     /tissue_type="venom glands"
                     /clone_lib="Snake Bothrops insularis library IL2"
                     /note="Organ: venom glands; Vector: pGEM11Zf+; Site_1: Eco
                     RI; Site_2: Not I; 5ug of mRNA from Bothrops insularis
                     venom glands were primed with oligo-(dT) and reverse
                     transcribed to cDNA using Superscript Plasmid System for
                     cDNA Synthesis and Cloning (LifeTechnologies). The cDNAs
                     were selected by size (350-600 pb and up 600 pb) in
                     agarose gel electrophoresis, linked to Eco RI adapters and
                     directionally cloned in pGEM11Zf+ vector (Promega). ESTs
                     were generated from random clones and grouped in unique
                     sequences. The putative identification of each EST or
                     cluster was obtained through Blast searches (e-value <
                     e-05)."
BASE COUNT      154 a    156 c    175 g    161 t      8 others
ORIGIN

Alignment Scores:
Pred. No.:              5.65e-22        Length:        654
Score:                  279.00          Matches:       61
Percent Similarity:     60.77%          Conservative:  18
Best Local Similarity:  46.92%          Mismatches:    43
Query Match:            39.02%          Indels:        8
DB:                     12              Gaps:          5

US-09-938-114-2 (1-129) x BM401401 (1-654)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      |||||||||||||||||||   ||||||:::||||||:::|||
Db        143 GATTGTCCCCCTGATTGGTCCTCCTATGAAGGGAGTTGCTACAGGGTCTTCGAACAAAAG 202

Qy         21 LysThrTrpThrAspAlaGluSerPheCysThrLysGlnValAsnGlyGlyHisLeuVal 40
                 |||   |||||||||   |||||||||:::|||      ||||||||||||||
Db        203 ATGAACTGGGAGGATGCAGAGAAATTCTGCACACAACAGCAGACAGGCGGACATCTGGTC 262

Qy         41 SerIleGluSerSerGlyGluAlaAspPheValGlyGlnLeuIleAlaGlnLysIleLys 60
              |||   :::||||||   ||||||||||||||       |||   :::      ::::::
Db        263 TCCTTCCAGAGCAGCGAAGAAGCAGATTTTGTGGTGTCGCTCACCTCACCAATTTTGAGA 322

Qy         61 SerAlaLysIleHisValTrpIleGlyLeuArgAlaGlnAsnLysGluLysGlnCysSer 80
              :::        ||||||   ||||||      :::      |||   |||
Db        323 GACAGT------TTTGTCTGGACCGGACTG------AGCGATGTCTGGAAAAGGATGCAGG 370

Qy         81 IleGluTrpSerAspGlySerSerIleSer---LysGluAsnTrpIleGluGluGluSer 99
              ||||||||||||||||||   :::      |||:::|||:::            |||
Db        371 TTCGAGTGGAGCGATGGCAGCGACCTCANCTACAAGGACAACTACCAGTTTGTGTTTTCA 430

Qy        100 Lys---LysCysLeuGlyValHisIleGluThrGlyPheHisLysTrpGluAsnPheTyr 118
              :::   :::|||      |||      :::|||      :::||||||
Db        431 GAATATGAGTGT------GTCGCATCAAAGACAAAAAATAACAAATGGAGGATTATACCC 484

Qy        119 CysGluGlnGlnAspProPheValCysGlu 128
              |||   :::   :::   ||||||||||||
Db        485 TGCACAAAGTTGGAATATTTCGTCTGCGAG 514


Search completed: December  8, 2003, 17:24:30
Job time : 3345.94 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:45:44 ; Search time 17.4 Seconds
(without alignments)
264.544 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	170	100.0	129	24	ABU08798	Deinagkistrodon ac
2	138	81.2	116	13	AAR24427	Sequence of the pl
3	135	79.4	30	16	AAR72232	Vipera palestinae
4	135	79.4	43	16	AAR72234	Vipera palestinae
5	135	79.4	127	16	AAR72236	Vipera palestinae
6	134	78.8	151	23	AAE20178	Pigmy rattlesnake
7	132	77.6	50	13	AAR23889	N-terminal sequenc
8	132	77.6	146	24	ABU08799	Deinagkistrodon ac
9	131	77.1	123	14	AAR38225	Sequence of polype
10	130	76.5	32	16	AAR72231	Vipera palestinae
11	130	76.5	52	16	AAR72233	Vipera palestinae
12	130	76.5	132	16	AAR72235	Vipera palestinae
13	128	75.3	120	23	ABB83143	Ahylysantinfarctas
14	127	74.7	30	22	AAM51540	Snake venom blood
15	127	74.7	146	22	AAM51544	Snake venom blood
16	127	74.7	152	23	AAE20179	Pigmy rattlesnake
17	124	72.9	122	23	ABB83142	Ahylysantinfarctas
18	123.5	72.6	28	13	AAR23886	N-terminal sequenc
19	120	70.6	131	14	AAR38222	Sequence of polype
20	119	70.0	30	22	AAM51541	Snake venom blood
21	119	70.0	125	14	AAR45157	Botrocetin beta su
22	119	70.0	145	23	AAO14521	Korean adder snake
23	119	70.0	152	22	AAM51543	Snake venom blood
24	117	68.8	133	14	AAR45156	Botrocetin alpha s
25	116	68.2	30	17	AAW04861	N-terminal peptide
26	116	68.2	30	17	AAW04862	N-terminal peptide
27	116	68.2	107	14	AAR38227	Sequence of polype
28	114	67.1	122	14	AAR38231	Sequence of polype
29	112	65.9	132	14	AAR38226	Sequence of polype
30	111	65.3	36	13	AAR23885	N-terminal sequenc
31	111	65.3	38	16	AAR71978	Snake venom derive
32	111	65.3	110	23	AAO20975	110-mer central lo
33	111	65.3	126	16	AAR71979	Snake venom derive
34	111	65.3	126	21	AAY85627	Snake venom derive
35	111	65.3	126	23	AAO20973	126-mer wild-type
36	111	65.3	126	23	AAO20981	M8-G4 D54A mutant
37	111	65.3	126	23	AAO20982	M9-G4 Y58A mutant
38	111	65.3	126	23	AAO20983	M10-G4 K61A mutant
39	111	65.3	126	23	AAO20984	M11-G4 E62A mutant
40	111	65.3	126	23	AAO20985	M12-G4 Y63A mutant
41	111	65.3	126	23	AAO20986	M13-G4 R66A mutant
42	111	65.3	126	23	AAO20987	M14-G4 Y67A mutant
43	111	65.3	126	23	AAO20988	M15-G4 R100A mutan
44	111	65.3	126	23	AAO20989	M16-G4 D101A mutan
45	111	65.3	126	23	AAO20990	M17-G4 R103A mutan

ALIGNMENTS

RESULT 1
ABU08798
ID ABU08798 standard; Protein; 129 AA.
XX
AC ABU08798;
XX
DT 02-JUN-2003 (first entry)
XX
DE Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
KW Antithrombosis; alpha chain; fibrin hydrolysis; blood clot; enzyme;
KW platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW thrombosis.
XX
OS Deinagkistrodon acutus.
XX
FH Key Location/Qualifiers
FT Region 1..9
FT /note= "Antithrombosis enzyme alpha chain N-terminus not
FT encoded by abx93668"
FT Region 1..29
FT /note= "Specifically claimed in claim 7"
FT Misc-difference 91
FT /note= "Encoded by TAT"
XX
PN US6489451-B1.
XX

PD 03-DEC-2002.
XX
PF 10-APR-1998; 98US-0058740.
XX
PR 10-APR-1997; 97US-043886P.
XX
PA (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI Li BX, Cheng X;
XX
DR WPI; 2003-352116/33.
DR N-PSDB; ABX93668.
XX
PT New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT preventing and treating vaso-occulusive and thromboembolic disorders,
PT including myocardial infarction, restenosis, cerebral thrombosis and
PT unstable angina -
XX
PS Claim 12; Column 11; 19pp; English.
XX
CC The invention relates to a new *Deinagkistrodon* actus anti-thrombosis
CC enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC prevents platelet aggregation. The anti-thrombosis enzyme was
CC administered to rabbits intravenously. Thrombosis was determined before
CC and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC following administration at 0.005 micro/kg and this activity was
CC increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC composition is useful for preventing and treating vaso-occulusive and
CC thromboembolic disorders, including myocardial infarction, restenosis,
CC angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC peripheral arterial occulsion, stroke. It is also useful for treating
CC atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC diseases. The present sequence represents the amino acid sequence of the
CC Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
SQ Sequence 129 AA;

Query Match 100.0%; Score 170; DB 24; Length 129;
Best Local Similarity 100.0%; Pred. No. 8.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
||||||||||||||||||||||||||||
Db 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

RESULT 2
AAR24427
ID AAR24427 standard; Protein; 116 AA.
XX
AC AAR24427;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1992 (first entry)
XX
DE Sequence of the platelet glycoprotein GPIb inhibitor beta chain
DE (CHH-B-beta).
XX
KW Platelet adherence inhibitor; platelet anti-adhesive;
KW antithrombotic agent; von Willebrand Factor;
KW platelet glycoprotein GPIb-IX complex.
XX
OS Crotalus horridus horridus.
XX
PN WO9208472-A1.
XX
PD 29-MAY-1992.
XX
PF 14-NOV-1991; 91WO-US08516.
XX
PR 16-NOV-1990; 90US-0614443.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Scarborough RM;
XX
DR WPI; 1992-199936/24.
XX
PT Platelet antiadhesive peptide(s) obtd. from snake venom - also
PT inhibit thrombus formation; for treatment of arteriosclerosis,
PT atherosclerosis, acute myocardial infarction, chronic unstable
PT angina, etc.
XX
PS Example; Fig 6; 5pp; English.
XX
CC The PAA was purified from a solution of snake venom. Analysis of
CC the peak inhibitory fractions by SDS-PAGE revealed 2 major proteins
CC migrating with a mol. wt. 23-28 kD. Both were able to inhibit
CC botrocetin and ristocetin induced platelet agglutination. They were
CC called CHH-A and CHH-B. The earlier eluting subunit (CHH-B-beta)
CC and later eluting subunit (CHH-B-alpha) were individually submitted
CC to N-terminal sequence analysis. The complete amino acid sequence
CC for the alpha and beta chains are given in AAR24426 and AAR24427.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 116 AA;

Query Match 81.2%; Score 138; DB 13; Length 116;
Best Local Similarity 75.9%; Pred. No. 2e-11;
Matches 22; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
|| |||||||||||:||:| || ||| |
Db 1 DCPSDWSSYEGHCYRVFQQEMTWDDAEKF 29

RESULT 3
AAR72232
ID AAR72232 standard; peptide; 30 AA.
XX
AC AAR72232;
XX
DT 25-MAR-2003 (updated)
DT 04-DEC-1995 (first entry)
XX
DE Vipera palestinae venom derived peptide N-terminal fragment.
XX
KW Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
KW platelet binding; von Willebrand factor.
XX
OS Vipera palestinae.
XX
PN WO9509183-A1.
XX
PD 06-APR-1995.
XX
PF 27-SEP-1994; 94WO-JP01583.
XX
PR 28-SEP-1993; 93JP-0241666.
PR 14-JAN-1994; 94JP-0002691.
PR 10-JUN-1994; 94JP-0128518.
XX
PA (AJIN) AJINOMOTO KK.
XX
PI Fukuchi N, Ishii K, Kaida K, Kobayashi T;
XX
DR WPI; 1995-147392/19.
XX
PT Peptide derived from Vipera palestinae venom - inhibits binding
PT of platelets to von Willebrand factor, useful as antithrombotic
PT agent
XX
PS Claim 3; Page 17; 37pp; Japanese.

```
XX
CC   AAR72232 is a Vipera palestinae venom derived N-terminal fragment,
CC   the full peptide is given in AAR72236. A peptide compsn. comprising
CC   AAR72236 and AAR72235 is claimed, it inhibits the binding of platelets
CC   to von Willebrand factors and can therefore be used as an anti-
CC   thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   30 AA;

  Query Match             79.4%;  Score 135;  DB 16;  Length 30;
  Best Local Similarity   75.9%;  Pred. No. 1.1e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLDKTWADAEKF 29


RESULT 4
AAR72234
ID   AAR72234 standard; peptide; 43 AA.
XX
AC   AAR72234;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide N-terminal fragment.
XX
KW   Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 4; Page 18; 37pp; Japanese.
XX
CC   AAR72234 is a Vipera palestinae venom derived N-terminal fragment,
CC   the full peptide is given in AAR72236. A peptide compsn. comprising
CC   AAR72236 and AAR72235 is claimed, it inhibits the binding of platelets
CC   to von Willebrand factors and can therefore be used as an anti-
CC   thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   43 AA;

  Query Match             79.4%;  Score 135;  DB 16;  Length 43;
  Best Local Similarity   75.9%;  Pred. No. 1.6e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLDKTWADAEKF 29


RESULT 5
AAR72236
ID   AAR72236 standard; peptide; 127 AA.
XX
AC   AAR72236;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide.
XX
KW   Vipera palestinae; venom; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 5; Pages 19-20; 37pp; Japanese.
XX
CC   AAR72236 is a Vipera palestinae venom derived peptide. A peptide
CC   compsn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
CC   binding of platelets to von Willebrand factors and can therefore
CC   be used as an anti- thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   127 AA;

  Query Match             79.4%;  Score 135;  DB 16;  Length 127;
  Best Local Similarity   75.9%;  Pred. No. 5.6e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLDKTWADAEKF 29


RESULT 6
AAE20178
ID   AAE20178 standard; Protein; 151 AA.
XX
AC   AAE20178;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk2.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; Zsnk2.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
```

```
FT   Peptide          1..19
FT                    /label= Signal_peptide
FT   Protein          20..151
FT                    /label= Mature_Zsnk2_protein
FT   Disulfide-bond   21..32
FT   Disulfide-bond   49..140
FT   Disulfide-bond   115..132
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   N-PSDB; AAD32053.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 3; Page 2; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk2.
XX
SQ   Sequence   151 AA;

  Query Match             78.8%;  Score 134;  DB 23;  Length 151;
  Best Local Similarity   75.9%;  Pred. No. 9.5e-11;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||: |||||| : ||| |||||
Db         20 DCPSDWSSYDQHCYKVFSELKTWDDAESF 48


RESULT 7
AAR23889
ID   AAR23889 standard; Protein; 50 AA.
XX
AC   AAR23889;
XX
DT   25-MAR-2003  (updated)
DT   21-NOV-1992  (first entry)
XX
DE   N-terminal sequence of the earlier eluting subunit PP-beta of
DE   the later eluting peak of the platelet antiadhesive peptide (PAA).
XX
KW   Platelet adherence inhibitor; platelet anti-adhesive;
KW   antithrombotic agent; von Willebrand Factor;
KW   platelet glycoprotein GPIb-IX complex.
XX
OS   Pseudocerastes persicus.
XX
PN   WO9208472-A1.
XX
PD   29-MAY-1992.
XX
PF   14-NOV-1991;   91WO-US08516.
XX
PR   16-NOV-1990;   90US-0614443.
XX
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Scarborough RM;
XX
DR   WPI; 1992-199936/24.
XX
PT   Platelet antiadhesive peptide(s) obtd. from snake venom - also
PT   inhibit thrombus formation; for treatment of arteriosclerosis,
PT   atherosclerosis, acute myocardial infarction, chronic unstable
PT   angina, etc.
XX
PS   Claim 7; Page 30; 5pp; English.
XX
CC   The PAA was purified from a solution of snake venom.  The earlier
CC   eluting subunit, PP-beta was subjected to Edman degradation for 50
CC   cylces to obtain the sequence in AAR23889.  The later eluting
CC   sequence - the PP-alpha chain was determined for 31 cycles to give
CC   the SQ in AAR23890.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   50 AA;

  Query Match             77.6%;  Score 132;  DB 13;  Length 50;
  Best Local Similarity   75.9%;  Pred. No. 5e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:|||||||    ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLYKTWEDAEKF 29


RESULT 8
ABU08799
ID   ABU08799 standard; Protein; 146 AA.
XX
AC   ABU08799;
XX
DT   02-JUN-2003  (first entry)
XX
DE   Deinagkistrodon acutus antithrombosis enzyme beta chain.
XX
KW   Antithrombosis; beta chain; fibrin hydrolysis; blood clot; enzyme;
KW   platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW   myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW   angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW   ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW   unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW   peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW   thrombosis.
XX
OS   Deinagkistrodon acutus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..23
FT                   /note= "Leader peptide"
FT   Protein         24..146
FT                   /note= "Mature antithrombosis enzyme beta chain. Residues
FT                           24 to 40 specifically claimed in claim 8"
```

```
FT   Misc-difference 72
FT                   /label= Unknown
FT                   /note= "Encoded by TNN"
FT   Misc-difference 72
FT                   /label= Unknown
FT                   /note= "Encoded by NNN"
FT   Misc-difference 72
FT                   /label= Unknown
FT                   /note= "Encoded by NNN"
FT   Misc-difference 72
FT                   /label= Unknown
FT                   /note= "Encoded by NNT"
XX
PN   US6489451-B1.
XX
PD   03-DEC-2002.
XX
PF   10-APR-1998;   98US-0058740.
XX
PR   10-APR-1997;   97US-043886P.
XX
PA   (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI   Li BX,  Cheng X;
XX
DR   WPI; 2003-352116/33.
DR   N-PSDB; ABX93674.
XX
PT   New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT   preventing and treating vaso-occulusive and thromboembolic disorders,
PT   including myocardial infarction, restenosis, cerebral thrombosis and
PT   unstable angina  -
XX
PS   Disclosure; Fig 1; 19pp; English.
XX
CC   The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC   enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC   prevents platelet aggregation. The anti-thrombosis enzyme was
CC   administered to rabbits intravenously. Thrombosis was determined before
CC   and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC   Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC   following administration at 0.005 micro/kg and this activity was
CC   increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC   composition is useful for preventing and treating vaso-occulusive and
CC   thromboembolic disorders, including myocardial infarction, restenosis,
CC   angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC   diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC   thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC   peripheral arterial occulsion, stroke. It is also useful for treating
CC   atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC   diseases. The present sequence represents the amino acid sequence of the
CC   Deinagkistrodon acutus antithrombosis enzyme beta chain.
XX
SQ   Sequence   146 AA;

  Query Match             77.6%;  Score 132;  DB 24;  Length 146;
  Best Local Similarity   72.4%;  Pred. No. 1.7e-10;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |:|||||||||| | : ||| ||| |
Db         24 DCPSEWSSYEGHCYKPFDEPKTWADAEKF 52


RESULT 9
AAR38225
ID   AAR38225 standard; protein; 123 AA.
XX
AC   AAR38225;
XX
DT   25-MAR-2003  (updated)
DT   01-OCT-1993  (first entry)
XX
DE   Sequence of polypeptide chain of alboaggregin A (AL-A).
XX
KW   Venom; snake; platelet-binding protein.
XX
OS   Trimeresurus albolabris.
XX
PN   WO9311151-A1.
XX
PD   10-JUN-1993.
XX
PF   01-DEC-1992;   92WO-US10344.
XX
PR   03-DEC-1991;   91US-0803630.
PR   05-JUN-1992;   92US-0893929.
XX
PA   (UTEM ) UNIV TEMPLE.
XX
PI   Kirby EP,  Peng M;
XX
DR   WPI; 1993-196991/24.
XX
PT   New platelet-binding proteins obtained from snake venom - inhibit
PT   binding of von Willebrand factor to platelet  membrane
PT   glycoprotein IB, useful in therapy, diagnosis and surgery
XX
PS   Claim 4; Page 47; 74pp; English.
XX
CC   AL-A may be obtd. from snake venom. It has a mol. wt. of about 45
CC   kDa. AL-A contains two types of polypeptide chains, with mol. wt. of
CC   about 18 kDa and about 15 kDa, respectively. Each of these two types
CC   of chains is actually composed of two subtypes. Thus, AL-A comprises
CC   four non-identical polypeptide chains. A possible variation of
CC   AAR38224 has been found, characterised by Asp at posn. three in lieu
CC   of Cys.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   123 AA;

  Query Match             77.1%;  Score 131;  DB 14;  Length 123;
  Best Local Similarity   69.0%;  Pred. No. 1.9e-10;
  Matches   20;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||:|| : : | ||| |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKF 29


RESULT 10
AAR72231
ID   AAR72231 standard; peptide; 32 AA.
XX
AC   AAR72231;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide N-terminal fragment.
XX
KW   Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
```

```
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 3; Page 17; 37pp; Japanese.
XX
CC   AAR72231 is a Vipera palestinae venom derived N-terminal fragment,
CC   the full peptide is given in AAR72235. A peptide compsn. comprising
CC   AAR72235 and AAR72236 is claimed, it inhibits the binding of platelets
CC   to von Willebrand factors and can therefore be used as an anti-
CC   thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   32 AA;

  Query Match             76.5%;  Score 130;  DB 16;  Length 32;
  Best Local Similarity   72.4%;  Pred. No. 5.7e-11;
  Matches   21;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:|||||||||:  ||| :|| |
Db          3 DCPSDWSSHEGHCYKVFRLFKTWEEAEKF 31


RESULT 11
AAR72233
ID   AAR72233 standard; peptide; 52 AA.
XX
AC   AAR72233;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide N-terminal fragment.
XX
KW   Vipera palestinae; venom; N-terminal fragment; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 4; Page 18; 37pp; Japanese.
XX
CC   AAR72233 is a Vipera palestinae venom derived N-terminal fragment,
CC   the full peptide is given in AAR72235. A peptide compsn. comprising
CC   AAR72235 and AAR72236 is claimed, it inhibits the binding of platelets
CC   to von Willebrand factors and can therefore be used as an anti-
CC   thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   52 AA;

  Query Match             76.5%;  Score 130;  DB 16;  Length 52;
  Best Local Similarity   72.4%;  Pred. No. 9.9e-11;
  Matches   21;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:|||||||||:  ||| :|| |
Db          3 DCPSDWSSHEGHCYKVFRLFKTWEEAEKF 31


RESULT 12
AAR72235
ID   AAR72235 standard; peptide; 132 AA.
XX
AC   AAR72235;
XX
DT   25-MAR-2003  (updated)
DT   04-DEC-1995  (first entry)
XX
DE   Vipera palestinae venom derived peptide.
XX
KW   Vipera palestinae; venom; antithrombotic agent;
KW   platelet binding; von Willebrand factor.
XX
OS   Vipera palestinae.
XX
PN   WO9509183-A1.
XX
PD   06-APR-1995.
XX
PF   27-SEP-1994;   94WO-JP01583.
XX
PR   28-SEP-1993;   93JP-0241666.
PR   14-JAN-1994;   94JP-0002691.
PR   10-JUN-1994;   94JP-0128518.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
PI   Fukuchi N,  Ishii K,  Kaida K,  Kobayashi T;
XX
DR   WPI; 1995-147392/19.
XX
PT   Peptide derived from Vipera palestinae venom - inhibits binding
PT   of platelets to von Willebrand factor, useful as antithrombotic
PT   agent
XX
PS   Claim 5; Page 19; 37pp; Japanese.
XX
CC   AAR72235 is a Vipera palestinae venom derived peptide. A peptide
CC   compsn. comprising AAR72235 and AAR72236 is claimed, it inhibits the
CC   binding of platelets to von Willebrand factors and can therefore
CC   be used as an anti- thrombotic agent.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   132 AA;

  Query Match             76.5%;  Score 130;  DB 16;  Length 132;
  Best Local Similarity   72.4%;  Pred. No. 2.9e-10;
  Matches   21;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:|||||||||:  ||| :|| |
Db          3 DCPSDWSSHEGHCYKVFRLFKTWEEAEKF 31


RESULT 13
ABB83143
ID   ABB83143 standard; protein; 120 AA.
XX
AC   ABB83143;
```

```
XX
DT   06-AUG-2002  (first entry)
XX
DE   Ahylysantinfarctase thrombase II subunit B.
XX
KW   Ahylysantinfarctase thrombase II subunit B; snake; venom;
KW   Chinese Agkistrodon acutus; purification.
XX
OS   Agkistrodon acutus.
XX
PN   CN1332242-A.
XX
PD   23-JAN-2002.
XX
PF   29-APR-2001; 2001CN-0115570.
XX
PR   29-APR-2001; 2001CN-0115570.
XX
PA   (KUNM-) KUNMING ZOOLOGY INST CHINESE ACAD SCI.
XX
PI   Xiao C;
XX
DR   WPI; 2002-281833/33.
XX
PT   Ahylysantinfarctase thrombase and its production process -
XX
PS   Claim 4; Page 5 (Disclosure); 8pp; Chinese.
XX
CC   The present invention relates to a method for purifying
CC   Ahylysantinfarctase thrombase II from the snake venom of Chinese
CC   Agkistrodon acutus. The method involves using an anionic exchange column
CC   chromatographic process to separate and purify and the repurifying the
CC   thrombase in a fast protein purifying work station. It was found that the
CC   Ahylysantinfarctase thrombase II consists of two subunits, Subunit A
CC   (ABB83142) and Subunit B (the present sequence).
XX
SQ   Sequence   120 AA;

  Query Match             75.3%;  Score 128;  DB 23;  Length 120;
  Best Local Similarity   69.0%;  Pred. No. 4.9e-10;
  Matches   20;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||:||||| |    | | |||:|
Db          1 DCPSDWSSYDGHCYKPFNDPKNWADAENF 29


RESULT 14
AAM51540
ID   AAM51540 standard; peptide; 30 AA.
XX
AC   AAM51540;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant peptide #1.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Example 1; Page 8; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence is
CC   a peptide derived from the protein.
XX
SQ   Sequence   30 AA;

  Query Match             74.7%;  Score 127;  DB 22;  Length 30;
  Best Local Similarity   69.0%;  Pred. No. 1.4e-10;
  Matches   20;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | |||||||||| | : | | |||:|
Db          1 DCPSGWSSYEGHCYKPFNEQKNWADAENF 29


RESULT 15
AAM51544
ID   AAM51544 standard; Protein; 146 AA.
XX
AC   AAM51544;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin B chain.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   N-PSDB; AAI71877.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 2; Page 11-12; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   is the B chain of halyxin.
XX
SQ   Sequence   146 AA;

  Query Match             74.7%;  Score 127;  DB 22;  Length 146;
  Best Local Similarity   69.0%;  Pred. No. 8.5e-10;
  Matches   20;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | |||||||||| | : | | |||:|
Db         24 DCPSGWSSYEGHCYKPFNEQKNWADAENF 52
```

Search completed: December 8, 2003, 09:52:34
Job time : 17.4 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:50:40 ; Search time 6.29714 Seconds
(without alignments)
194.853 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	170	100.0	29	4	US-09-058-740-3	Sequence 3, Appli
2	170	100.0	129	4	US-09-058-740-2	Sequence 2, Appli
3	138	81.2	117	1	US-07-614-443A-2	Sequence 2, Appli
4	138	81.2	117	1	US-08-294-859-2	Sequence 2, Appli
5	138	81.2	117	1	US-08-481-676-2	Sequence 2, Appli
6	132	77.6	50	1	US-07-614-443A-5	Sequence 5, Appli
7	132	77.6	50	1	US-08-294-859-5	Sequence 5, Appli
8	132	77.6	50	1	US-08-481-676-5	Sequence 5, Appli
9	131	77.1	123	1	US-07-893-929A-4	Sequence 4, Appli
10	131	77.1	123	5	PCT-US92-10344-4	Sequence 4, Appli
11	120	70.6	131	1	US-07-893-929A-1	Sequence 1, Appli
12	120	70.6	131	5	PCT-US92-10344-1	Sequence 1, Appli
13	116	68.2	30	2	US-08-894-403-3	Sequence 3, Appli
14	116	68.2	30	2	US-08-894-403-4	Sequence 4, Appli
15	116	68.2	107	1	US-07-893-929A-6	Sequence 6, Appli
16	116	68.2	107	5	PCT-US92-10344-6	Sequence 6, Appli
17	114	67.1	123	1	US-07-893-929A-10	Sequence 10, Appl
18	114	67.1	123	5	PCT-US92-10344-10	Sequence 10, Appl
19	113	66.5	128	1	US-07-893-929A-8	Sequence 8, Appli
20	113	66.5	128	5	PCT-US92-10344-8	Sequence 8, Appli
21	112	65.9	132	1	US-07-893-929A-5	Sequence 5, Appli
22	112	65.9	132	5	PCT-US92-10344-5	Sequence 5, Appli
23	111	65.3	38	2	US-08-612-840A-1	Sequence 1, Appli
24	111	65.3	126	2	US-08-612-840A-2	Sequence 2, Appli
25	111	65.3	127	1	US-07-614-443A-1	Sequence 1, Appli
26	111	65.3	127	1	US-08-294-859-1	Sequence 1, Appli
27	111	65.3	127	1	US-08-481-676-1	Sequence 1, Appli
28	111	65.3	130	1	US-07-893-929A-7	Sequence 7, Appli
29	111	65.3	130	5	PCT-US92-10344-7	Sequence 7, Appli
30	111	65.3	149	2	US-08-612-840A-8	Sequence 8, Appli
31	107	62.9	125	1	US-07-893-929A-3	Sequence 3, Appli
32	107	62.9	125	5	PCT-US92-10344-3	Sequence 3, Appli
33	104	61.2	31	1	US-07-614-443A-6	Sequence 6, Appli
34	104	61.2	31	1	US-08-294-859-6	Sequence 6, Appli
35	104	61.2	31	1	US-08-481-676-6	Sequence 6, Appli
36	103	60.6	21	2	US-08-612-840A-3	Sequence 3, Appli
37	102.5	60.3	133	1	US-07-893-929A-9	Sequence 9, Appli
38	102.5	60.3	133	5	PCT-US92-10344-9	Sequence 9, Appli
39	94	55.3	134	1	US-07-893-929A-2	Sequence 2, Appli
40	94	55.3	134	5	PCT-US92-10344-2	Sequence 2, Appli
41	84	49.4	1257	1	US-08-340-428B-49	Sequence 49, Appl
42	82	48.2	908	5	PCT-US95-03747-3	Sequence 3, Appli
43	81	47.6	17	4	US-09-058-740-4	Sequence 4, Appli
44	79	46.5	912	5	PCT-US95-03747-2	Sequence 2, Appli
45	75	44.1	2409	6	5180808-2	Patent No. 5180808

ALIGNMENTS

```
RESULT 1
US-09-058-740-3
; Sequence 3, Application US/09058740
; Patent No. 6489451
;  GENERAL INFORMATION:
;       APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                  Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                  Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                  Jin-Guo Ding, Fang Rong, Yan Liu and
;                  Hui-Ran Chen
;       TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                           VENOM OF AGKISTRODON ACUTUS
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/058,740
;            FILING DATE: 10-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: <Unknown>
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Chen, Anthony C.
;            REGISTRATION NUMBER: 38,673
;            REFERENCE/DOCKET NUMBER: 233/298
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 29 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: Peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
```

```
US-09-058-740-3

  Query Match             100.0%;  Score 170;  DB 4;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-16;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29


RESULT 2
US-09-058-740-2
; Sequence 2, Application US/09058740
; Patent No. 6489451
;   GENERAL INFORMATION:
;        APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                   Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                   Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                   Jin-Guo Ding, Fang Rong, Yan Liu and
;                   Hui-Ran Chen
;        TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                            VENOM OF AGKISTRODON ACUTUS
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/058,740
;             FILING DATE: 10-Apr-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Chen, Anthony C.
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 129 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: Peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-058-740-2

  Query Match             100.0%;  Score 170;  DB 4;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-15;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29


RESULT 3
```

```
US-07-614-443A-2
; Sequence 2, Application US/07614443A
; Patent No. 5342830
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/614,443A
;      FILING DATE:  19901116
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22803-20003.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-614-443A-2

  Query Match             81.2%;  Score 138;  DB 1;  Length 117;
  Best Local Similarity   75.9%;  Pred. No. 2e-11;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||:||:|  || ||| |
Db          1 DCPSDWSSYEGHCYRVFQQEMTWDDAEKF 29


RESULT 4
US-08-294-859-2
; Sequence 2, Application US/08294859
; Patent No. 5679542
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/294,859
;      FILING DATE:  29-AUG-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
```

```
;       REGISTRATION NUMBER:  29,959
;       REFERENCE/DOCKET NUMBER:  2803-0003.10
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 887-1500
;       TELEFAX:  (202) 887-0763
;       TELEX:  90-4030
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  117 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-294-859-2

  Query Match             81.2%;  Score 138;  DB 1;  Length 117;
  Best Local Similarity   75.9%;  Pred. No. 2e-11;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||:||:|  || ||| |
Db          1 DCPSDWSSYEGHCYRVFQQEMTWDDAEKF 29


RESULT 5
US-08-481-676-2
; Sequence 2, Application US/08481676
; Patent No. 5744584
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,676
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/294,859
;      FILING DATE:  29-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2803-0003.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  117 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-481-676-2

  Query Match             81.2%;  Score 138;  DB 1;  Length 117;
  Best Local Similarity   75.9%;  Pred. No. 2e-11;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||:||:|  || ||| |
Db          1 DCPSDWSSYEGHCYRVFQQEMTWDDAEKF 29


RESULT 6
US-07-614-443A-5
; Sequence 5, Application US/07614443A
; Patent No. 5342830
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  755 Page Mill Road
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/614,443A
;      FILING DATE:  19901116
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  22803-20003.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-614-443A-5

  Query Match             77.6%;  Score 132;  DB 1;  Length 50;
  Best Local Similarity   75.9%;  Pred. No. 5e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLYKTWEDAEKF 29


RESULT 7
US-08-294-859-5
; Sequence 5, Application US/08294859
; Patent No. 5679542
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/294,859
;      FILING DATE:  29-AUG-1994
;      CLASSIFICATION:  435
```

```
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Murashige, Kate H.
;       REGISTRATION NUMBER:  29,959
;       REFERENCE/DOCKET NUMBER:  2803-0003.10
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 887-1500
;       TELEFAX:  (202) 887-0763
;       TELEX:  90-4030
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  50 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-294-859-5

  Query Match             77.6%;  Score 132;  DB 1;  Length 50;
  Best Local Similarity   75.9%;  Pred. No. 5e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLYKTWEDAEKF 29


RESULT 8
US-08-481-676-5
; Sequence 5, Application US/08481676
; Patent No. 5744584
;  GENERAL INFORMATION:
;    APPLICANT:  SCARBOROUGH, ROBERT M.
;    TITLE OF INVENTION:  ANTITHROMBOSIS AGENTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Ave. N.W., Suite 5500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,676
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/294,859
;      FILING DATE:  29-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  2803-0003.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-481-676-5

  Query Match             77.6%;  Score 132;  DB 1;  Length 50;
  Best Local Similarity   75.9%;  Pred. No. 5e-11;
  Matches   22;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:||||||||   ||| ||| |
Db          1 DCPSDWSSHEGHCYKVFNLYKTWEDAEKF 29


RESULT 9
US-07-893-929A-4
; Sequence 4, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  123 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-893-929A-4

  Query Match             77.1%;  Score 131;  DB 1;  Length 123;
  Best Local Similarity   69.0%;  Pred. No. 1.8e-10;
  Matches   20;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||:|| : : | ||| |
Db          1 DCPSDWSSYEGHCYRVFNEPQNWADAEKF 29


RESULT 10
PCT-US92-10344-4
; Sequence 4, Application PC/TUS9210344
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
```

; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-10344-4

Query Match 77.1%; Score 131; DB 5; Length 123;
Best Local Similarity 69.0%; Pred. No. 1.8e-10;
Matches 20; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
|| |||||||||||:|| : : | ||| |
Db 1 DCPSDWSSYEGHCYRVFNEPQNWADAEKF 29

RESULT 11
US-07-893-929A-1
; Sequence 1, Application US/07893929A
; Patent No. 5336667
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/893,929A
; FILING DATE: 19920605
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5336667e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-893-929A-1

Query Match 70.6%; Score 120; DB 1; Length 131;
Best Local Similarity 65.5%; Pred. No. 5.7e-09;
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
|| |||||||: :||:|||: :|| ||| |
Db 1 DCPSDWSSYDQYCYRVFKRLQTWEDAERF 29

RESULT 12
PCT-US92-10344-1
; Sequence 1, Application PC/TUS9210344
; GENERAL INFORMATION:
; APPLICANT: Kirby, Edward P.
; APPLICANT: Peng, Man-ling
; TITLE OF INVENTION: Alboaggregins: Platelet
; TITLE OF INVENTION: Agonists Which Bind To Platelet
; TITLE OF INVENTION: Membrane Glycoprotein Ib
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Temple University - Of The Common-
; ADDRESSEE: wealth System of Higher Education
; STREET: 406 University Services Building
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10344
; FILING DATE: 19921201
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/803,630
; FILING DATE: December 3, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 6056-126 (CIP) 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: None
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-10344-1

Query Match 70.6%; Score 120; DB 5; Length 131;
Best Local Similarity 65.5%; Pred. No. 5.7e-09;
Matches 19; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
|| |||||||: :||:|||: :|| ||| |
Db 1 DCPSDWSSYDQYCYRVFKRLQTWEDAERF 29

```
RESULT 13
US-08-894-403-3
; Sequence 3, Application US/08894403A
; Patent No. 5972681
; Patent No. 5972681 5972861
; GENERAL INFORMATION:
;  APPLICANT: MORITA, Takashi
;  TITLE OF INVENTION: CALCIUM-REQUIRING PROTHROMBIN ACTIVATOR
;  FILE REFERENCE: 425-653P
;  CURRENT APPLICATION NUMBER: US/08/894,403A
;  CURRENT FILING DATE: 1997-08-18
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-08-894-403-3

  Query Match             68.2%;  Score 116;  DB 2;  Length 30;
  Best Local Similarity   65.5%;  Pred. No. 4e-09;
  Matches   19;  Conservative    1;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              ||   |||:|||||||| |   | ||| |
Db          1 DCLPGWSSHEGHCYKVFNQEMYWADAEKF 29


RESULT 14
US-08-894-403-4
; Sequence 4, Application US/08894403A
; Patent No. 5972681
; Patent No. 5972681 5972861
; GENERAL INFORMATION:
;  APPLICANT: MORITA, Takashi
;  TITLE OF INVENTION: CALCIUM-REQUIRING PROTHROMBIN ACTIVATOR
;  FILE REFERENCE: 425-653P
;  CURRENT APPLICATION NUMBER: US/08/894,403A
;  CURRENT FILING DATE: 1997-08-18
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-08-894-403-4

  Query Match             68.2%;  Score 116;  DB 2;  Length 30;
  Best Local Similarity   62.1%;  Pred. No. 4e-09;
  Matches   18;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              ||  ||  ||||||:|| : | | ||| |
Db          1 DCLPDWFHYEGHCYRVFDEPKKWADAEKF 29


RESULT 15
US-07-893-929A-6
; Sequence 6, Application US/07893929A
; Patent No. 5336667
;  GENERAL INFORMATION:
;    APPLICANT:  Kirby, Edward P.
;    APPLICANT:  Peng, Man-ling
;    TITLE OF INVENTION:  Alboaggregins: Platelet
;    TITLE OF INVENTION:  Agonists Which Bind To Platelet
;    TITLE OF INVENTION:  Membrane Glycoprotein Ib
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Temple University - Of The Common-
;      ADDRESSEE:  wealth System of Higher Education
;      STREET:  406 University Services Building
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/893,929A
;      FILING DATE:  19920605
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/803,630
;      FILING DATE:  December 3, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  6056-126 (CIP) 1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5336667e
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  107 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
US-07-893-929A-6

  Query Match             68.2%;  Score 116;  DB 1;  Length 107;
  Best Local Similarity   65.5%;  Pred. No. 1.6e-08;
  Matches   19;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||: :|||||:|   | ||| |
Db          1 DCPSDWSSYDLYCYKVFQQRMNWEDAEQF 29


Search completed: December  8, 2003, 09:56:25
Job time : 6.29714 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:54:55 ; Search time 11.9314 Seconds
(without alignments)
452.044 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	170	100.0	29	11	US-09-938-114-3	Sequence 3, Appli
2	170	100.0	129	11	US-09-938-114-2	Sequence 2, Appli
3	134	78.8	151	10	US-09-929-230-2	Sequence 2, Appli
4	134	78.8	151	12	US-10-226-420-2	Sequence 2, Appli
5	127	74.7	152	10	US-09-929-230-5	Sequence 5, Appli
6	127	74.7	152	12	US-10-226-420-5	Sequence 5, Appli
7	111	65.3	149	10	US-09-969-763-3	Sequence 3, Appli
8	110	64.7	144	10	US-09-929-230-8	Sequence 8, Appli
9	110	64.7	144	12	US-10-226-420-8	Sequence 8, Appli
10	100	58.8	110	10	US-09-969-763-10	Sequence 10, Appl
11	97	57.1	158	10	US-09-929-230-11	Sequence 11, Appl
12	97	57.1	158	12	US-10-226-420-11	Sequence 11, Appl
13	96	56.5	126	10	US-09-969-763-1	Sequence 1, Appli
14	84	49.4	52	12	US-10-029-386-31252	Sequence 31252, A
15	84	49.4	1321	12	US-10-241-220-82	Sequence 82, Appl
16	82	48.2	136	10	US-09-965-528-24	Sequence 24, Appl
17	82	48.2	164	9	US-09-801-438-2	Sequence 2, Appli
18	82	48.2	164	10	US-09-893-737-20	Sequence 20, Appl
19	82	48.2	883	11	US-09-759-130B-342	Sequence 342, App
20	82	48.2	883	12	US-10-188-495-72	Sequence 72, Appl
21	82	48.2	883	15	US-10-195-970-3	Sequence 3, Appli
22	82	48.2	883	15	US-10-195-970-6	Sequence 6, Appli
23	82	48.2	883	15	US-10-189-123-72	Sequence 72, Appl
24	81	47.6	17	11	US-09-938-114-4	Sequence 4, Appli
25	77	45.3	29	10	US-09-929-230-14	Sequence 14, Appl
26	77	45.3	29	12	US-10-226-420-14	Sequence 14, Appl
27	76	44.7	211	15	US-10-050-704-266	Sequence 266, App
28	76	44.7	427	15	US-10-050-704-273	Sequence 273, App
29	76	44.7	483	15	US-10-050-704-272	Sequence 272, App
30	76	44.7	911	12	US-10-237-496-52	Sequence 52, Appl
31	76	44.7	911	12	US-10-242-074-52	Sequence 52, Appl
32	76	44.7	911	12	US-10-242-505-52	Sequence 52, Appl
33	76	44.7	911	12	US-10-242-574-52	Sequence 52, Appl
34	76	44.7	911	12	US-10-243-261-52	Sequence 52, Appl
35	76	44.7	911	12	US-10-243-282-52	Sequence 52, Appl
36	76	44.7	911	12	US-10-243-402-52	Sequence 52, Appl
37	76	44.7	911	12	US-10-243-431-52	Sequence 52, Appl
38	76	44.7	911	12	US-10-245-164-52	Sequence 52, Appl
39	76	44.7	911	12	US-10-244-972-52	Sequence 52, Appl
40	76	44.7	911	12	US-10-197-942-52	Sequence 52, Appl
41	76	44.7	911	12	US-10-238-196-52	Sequence 52, Appl
42	76	44.7	911	12	US-10-245-013-52	Sequence 52, Appl
43	76	44.7	911	15	US-10-245-103-52	Sequence 52, Appl
44	76	44.7	911	15	US-10-245-107-52	Sequence 52, Appl
45	76	44.7	911	15	US-10-245-143-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-09-938-114-3
; Sequence 3, Application US/09938114
; Publication No. US20030022350A1
; GENERAL INFORMATION:
; APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
; Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
; Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
; Jin-Guo Ding, Fang Rong, Yan Liu and
; Hui-Ran Chen
; TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
; VENOM OF AGKISTRODON ACUTUS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/938,114
; FILING DATE: 23-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,740
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Chen, Anthony C.
; REGISTRATION NUMBER: 38,673
; REFERENCE/DOCKET NUMBER: 233/298

```
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (213) 489-1600
;              TELEFAX: (213) 955-0440
;              TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 3:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 29 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-938-114-3

  Query Match             100.0%;  Score 170;  DB 11;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-16;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29


RESULT 2
US-09-938-114-2
; Sequence 2, Application US/09938114
; Publication No. US20030022350A1
;   GENERAL INFORMATION:
;         APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                    Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                    Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                    Jin-Guo Ding, Fang Rong, Yan Liu and
;                    Hui-Ran Chen
;         TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                             VENOM OF AGKISTRODON ACUTUS
;         NUMBER OF SEQUENCES: 12
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Lyon & Lyon
;              STREET: 633 West Fifth Street
;                      Suite 4700
;              CITY: Los Angeles
;              STATE: California
;              COUNTRY: U.S.A.
;              ZIP: 90071-2066
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                           storage
;              COMPUTER: IBM Compatible
;              OPERATING SYSTEM: IBM P.C. DOS 5.0
;              SOFTWARE: FastSEQ for Windows 2.0
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/938,114
;              FILING DATE: 23-Aug-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/058,740
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Chen, Anthony C.
;              REGISTRATION NUMBER: 38,673
;              REFERENCE/DOCKET NUMBER: 233/298
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (213) 489-1600
;              TELEFAX: (213) 955-0440
;              TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 2:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 129 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: Peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-938-114-2

  Query Match             100.0%;  Score 170;  DB 11;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-15;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |||||||||||||||||||||||||||||
Db          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29


RESULT 3
US-09-929-230-2
; Sequence 2, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-2

  Query Match             78.8%;  Score 134;  DB 10;  Length 151;
  Best Local Similarity   75.9%;  Pred. No. 1.6e-10;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||| |||||| : ||| |||||
Db         20 DCPSDWSSYDQHCYKVFSELKTWDDAESF 48


RESULT 4
US-10-226-420-2
; Sequence 2, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-2

  Query Match             78.8%;  Score 134;  DB 12;  Length 151;
  Best Local Similarity   75.9%;  Pred. No. 1.6e-10;
  Matches   22;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||| |||||| : ||| |||||
Db         20 DCPSDWSSYDQHCYKVFSELKTWDDAESF 48


RESULT 5
US-09-929-230-5
; Sequence 5, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
```

```
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-5

  Query Match             74.7%;  Score 127;  DB 10;  Length 152;
  Best Local Similarity   72.4%;  Pred. No. 1.4e-09;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | ||||: |||:|||| ||| ||| |
Db         24 DCPSGWSSYDQHCYRVFKQLKTWDDAERF 52


RESULT 6
US-10-226-420-5
; Sequence 5, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-5

  Query Match             74.7%;  Score 127;  DB 12;  Length 152;
  Best Local Similarity   72.4%;  Pred. No. 1.4e-09;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | ||||: |||:|||| ||| ||| |
Db         24 DCPSGWSSYDQHCYRVFKQLKTWDDAERF 52


RESULT 7
US-09-969-763-3
; Sequence 3, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 149
;   TYPE: PRT
;   ORGANISM: Crotalus harridus
US-09-969-763-3

  Query Match             65.3%;  Score 111;  DB 10;  Length 149;
  Best Local Similarity   62.1%;  Pred. No. 2e-07;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :| | ||||: :||| |||  || ||| |
Db         26 ECPSGWSSYDRYCYKPFKQEMTWADAERF 54


RESULT 8
US-09-929-230-8
; Sequence 8, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 144
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-8

  Query Match             64.7%;  Score 110;  DB 10;  Length 144;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-07;
  Matches   17;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || ||| :|: :||:| || :|| ||| |
Db         12 DCPSDWYAYDQYCYRVIKQLRTWDDAERF 40


RESULT 9
US-10-226-420-8
; Sequence 8, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
;   LENGTH: 144
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-8

  Query Match             64.7%;  Score 110;  DB 12;  Length 144;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-07;
  Matches   17;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || ||| :|: :||:| || :|| ||| |
Db         12 DCPSDWYAYDQYCYRVIKQLRTWDDAERF 40


RESULT 10
US-09-969-763-10
; Sequence 10, Application US/09969763
```

```
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 110
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: SYNTHETIC PEPTIDE
US-09-969-763-10

  Query Match             58.8%;  Score 100;  DB 10;  Length 110;
  Best Local Similarity   58.6%;  Pred. No. 4.5e-06;
  Matches   17;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :| | ||||  :||| |||  |: ||| |
Db          3 ECPSGWSSYSRYCYKPFKQEMTYADAERF 31


RESULT 11
US-09-929-230-11
; Sequence 11, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-09-929-230-11

  Query Match             57.1%;  Score 97;  DB 10;  Length 158;
  Best Local Similarity   51.7%;  Pred. No. 1.6e-05;
  Matches   15;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :| | | :|: :||:| |: ||| ||| |
Db         26 NCPSGWFAYDQYCYRVIKRLKTWDDAERF 54


RESULT 12
US-10-226-420-11
; Sequence 11, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Sistrurus miliarius
US-10-226-420-11

  Query Match             57.1%;  Score 97;  DB 12;  Length 158;
  Best Local Similarity   51.7%;  Pred. No. 1.6e-05;
  Matches   15;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :| | | :|: :||:| |: ||| ||| |
Db         26 NCPSGWFAYDQYCYRVIKRLKTWDDAERF 54


RESULT 13
US-09-969-763-1
; Sequence 1, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING TH
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 126
;   TYPE: PRT
;   ORGANISM: Crotallus horridus
US-09-969-763-1

  Query Match             56.5%;  Score 96;  DB 10;  Length 126;
  Best Local Similarity   55.2%;  Pred. No. 1.8e-05;
  Matches   16;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :| | ||| : :||| |||  ||  || |
Db          3 ECPSGWSSTDRYCYKPFKQEMTWASAERF 31


RESULT 14
US-10-029-386-31252
; Sequence 31252, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31252
;   LENGTH: 52
```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005254.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
; OTHER INFORMATION: SWISSPROT HIT: P55067, EVALUE 7.00e-28
US-10-029-386-31252

  Query Match 49.4%; Score 84; DB 12; Length 52;
  Best Local Similarity 42.3%; Pred. No. 0.0003;
  Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CSSDWSSYEGHCYKVFKQSKTWTDAE 27
Db 4 CDRGWHKFQGHCYRYFAHRRAWEDAE 29

RESULT 15
US-10-241-220-82
; Sequence 82, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz,Gretchen
; APPLICANT: Hillan,Kenneth J.
; APPLICANT: Phillips,Heidi
; APPLICANT: Polakis,Paul
; APPLICANT: Spencer,Susan
; APPLICANT: Williams,P.Mickey
; APPLICANT: Wu,Thomas
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 82
; LENGTH: 1321
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-82

  Query Match 49.4%; Score 84; DB 12; Length 1321;
  Best Local Similarity 42.3%; Pred. No. 0.0082;
  Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CSSDWSSYEGHCYKVFKQSKTWTDAE 27
Db 1088 CDRGWHKFQGHCYRYFAHRRAWEDAE 1113

Search completed: December 8, 2003, 10:07:06
Job time : 11.9314 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:41:24 ; Search time 1230.93 Seconds
(without alignments)
963.811 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22849/app_query.fasta_1.725
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_3707_@runat_08122003_091000_22849 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	163	95.9	733	5	AF176420	AF176420 Deinagkis
2	155	91.2	497	5	AY091758	AY091758 Deinagkis
3	155	91.2	656	5	AB036880	AB036880 Deinagkis
4	148	87.1	688	5	AF540647	AF540647 Deinagkis
5	136	80.0	458	5	AF176421	AF176421 Deinagkis
6	134	78.8	483	6	AX427201	AX427201 Sequence
7	134	78.8	664	5	AB019616	AB019616 Agkistrod
8	133	78.2	478	5	AY091761	AY091761 Deinagkis
9	133	78.2	592	5	AB036881	AB036881 Deinagkis
10	133	78.2	698	5	TFLFIXB	D83332 Trimeresuru
11	132	77.6	369	5	AY293866	AY293866 Deinagkis
12	132	77.6	454	6	AR259041	AR259041 Sequence
13	132	77.6	478	5	AY091756	AY091756 Deinagkis
14	132	77.6	574	5	AF350324	AF350324 Deinagkis
15	127	74.7	634	5	AF197915	AF197915 Gloydius
16	127	74.7	721	6	AX427204	AX427204 Sequence
17	124	72.9	438	5	AF244901	AF244901 Callosela
18	124	72.9	484	5	AY091760	AY091760 Deinagkis
19	123	72.4	678	5	TFLFIXA	D83331 Trimeresuru
20	123	72.4	683	5	AB046491	AB046491 Trimeresu
21	121	71.2	375	5	AF387100	AF387100 Deinagkis
22	119	70.0	453	6	AX427203	AX427203 Sequence
23	119	70.0	583	5	AF125310	AF125310 Gloydius
24	119	70.0	630	5	AF190827	AF190827 Gloydius
25	116	68.2	478	5	AY091755	AY091755 Deinagkis
26	116	68.2	685	5	AF102902	AF102902 Deinagkis
27	116	68.2	738	5	CDRNACVXB	Y16349 Crotalus du
28	115	67.6	456	6	AX427206	AX427206 Sequence
29	115	67.6	544	6	AR259035	AR259035 Sequence
30	114	67.1	487	5	AF463521	AF463521 Deinagkis
31	114	67.1	565	5	AY268948	AY268948 Echis car
32	114	67.1	619	5	AF540648	AF540648 Deinagkis
33	111	65.3	690	6	AR026653	AR026653 Sequence
34	111	65.3	690	6	AX404807	AX404807 Sequence
35	111	65.3	690	6	BD143687	BD143687 Protein h
36	110	64.7	580	6	AX427207	AX427207 Sequence
37	103	60.6	408	5	AF244900	AF244900 Callosela
38	103	60.6	432	6	AX427209	AX427209 Sequence
39	103	60.6	508	5	AY091759	AY091759 Deinagkis
40	103	60.6	666	5	AF540645	AF540645 Deinagkis
41	102	60.0	441	5	AY099322	AY099322 Protoboth
42	102	60.0	681	5	AY149340	AY149340 Trimeresu
43	100	58.8	477	5	AY099321	AY099321 Protoboth
44	99	58.2	632	5	AF125309	AF125309 Gloydius
45	99	58.2	679	5	AB019615	AB019615 Agkistrod

ALIGNMENTS

RESULT 1

```
AF176420
LOCUS       AF176420                733 bp    mRNA    linear   VRT 30-APR-2002
DEFINITION  Deinagkistrodon acutus clone 2100490 agkisacutacin A chain mRNA,
            complete cds.
ACCESSION   AF176420 AY091757
VERSION     AF176420.2  GI:20273043
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J. and Liu,J.
  TITLE     cDNA sequencing and analysis of eleven C-type lectin-like protein
            subunits from Agkistrodon acutus
  JOURNAL   Sheng wu hua hsueh yu sheng wu wu li hsueh pao (2002) In press
REFERENCE   2  (bases 1 to 733)
  AUTHORS   Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and
            Liu,J.
  TITLE     Agkisacutacin, a new fibrinolytic & anti-platelet protein from
            Agkistrodon acutus venom
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     A chain of agkisacutacin from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 733)
  AUTHORS   Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
REFERENCE   5  (bases 1 to 733)
  AUTHORS   Yu,H.-X., Xiang,K.-J., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life-Science, Huangshan Road, Hefei, Anhui 230027, China
  REMARK    Sequence update by submitter
COMMENT     On Apr 23, 2002 this sequence version replaced gi:6715112.
FEATURES             Location/Qualifiers
     source          1. .733
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100490"
                     /tissue_type="venom"
     CDS             89. .547
                     /note="consists of two heterologous subunits (A and B
                     chains) linked by an inter-subunit disulfide bond; c-type
                     lectin family member"
                     /codon_start=1
                     /product="agkisacutacin A chain"
                     /protein_id="AAF26286.2"
                     /db_xref="GI:20273044"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCSSGWSSYEGHCYKVFKQSK
                     TWTDAESFCTKQVNGGHLVSIESSGEADFVGQLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHIETGFHKWENFYCEQQDPFVCEA"
BASE COUNT      191 a    173 c    204 g    165 t
ORIGIN

Alignment Scores:
Pred. No.:                  2.57e-16        Length:         733
Score:                      163.00          Matches:        28
Percent Similarity:         96.55%          Conservative:   0
Best Local Similarity:      96.55%          Mismatches:     1
Query Match:                95.88%          Indels:         0
DB:                         5               Gaps:           0

US-09-938-114-3 (1-29) x AF176420 (1-733)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        158 GATTGTTCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAGGTCTTCAAACAATCT 217

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||||||||||||||||||||
Db        218 AAGACCTGGACAGATGCAGAGAGCTTC 244

RESULT 2
AY091758
LOCUS       AY091758                497 bp    mRNA    linear   VRT 13-MAY-2002
DEFINITION  Deinagkistrodon acutus clone 2100590 ACF 1/2 A-chain mRNA, complete
            cds.
ACCESSION   AY091758
VERSION     AY091758.1  GI:20562936
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 497)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     A chain of ACF 1/2 from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 497)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES             Location/Qualifiers
     source          1. .497
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100590"
     CDS             1. .459
                     /note="C-type lectin family member"
                     /codon_start=1
                     /product="ACF 1/2 A-chain"
                     /protein_id="AAM22786.1"
                     /db_xref="GI:20562937"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCSSSWSSYEGHCYKAFKQSK
                     TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
BASE COUNT      135 a    102 c    143 g    117 t
ORIGIN

Alignment Scores:
Pred. No.:                  3.12e-15        Length:         497
Score:                      155.00          Matches:        26
Percent Similarity:         89.66%          Conservative:   0
Best Local Similarity:      89.66%          Mismatches:     3
Query Match:                91.18%          Indels:         0
DB:                         5               Gaps:           0

US-09-938-114-3 (1-29) x AY091758 (1-497)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         70 GATTGTTCCTCTAGTTGGTCCTCCTATGAAGGGCATTGCTACAAGGCCTTCAAACAATCT 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||  ||||||||||||||||
Db        130 AAGACCTGGGCAGATGCAGAGAGCTTC 156

RESULT 3
AB036880
LOCUS       AB036880                656 bp    mRNA    linear   VRT 23-MAY-2002
DEFINITION  Deinagkistrodon acutus acp-a mRNA for anticoagulant protein A,
            complete cds.
ACCESSION   AB036880
VERSION     AB036880.1  GI:8980618
```

```
KEYWORDS    anticoagulant protein A.
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1
  AUTHORS   Tani,A., Ogawa,T., Nose,T., Nikandrov,N.N., Deshimaru,M.,
            Chijiwa,T., Chang,C.C., Fukumaki,Y. and Ohno,M.
  TITLE     Characterization, primary structure and molecular evolution of
            anticoagulant protein from Agkistrodon actus venom
  JOURNAL   Toxicon 40 (6), 803-813 (2002)
  MEDLINE   21975857
REFERENCE   2  (bases 1 to 656)
  AUTHORS   Ogawa,T. and Tani,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate
            School of Agricultural Science; 1-1, Tsutusmidori Amamiyamachi,
            Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
            URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808,
            Fax:81-22-717-8807)
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /tissue_type="venom gland"
     gene            1. .656
                     /gene="acp-a"
     CDS             50. .508
                     /gene="acp-a"
                     /codon_start=1
                     /product="anticoagulant protein A"
                     /protein_id="BAA99281.1"
                     /db_xref="GI:8980619"
                     /translation="MGRFIFVSFGLLVVYLSLSGTAADCSSSWSSYEGHCYKAFKQSK
                     TWADAESFCTKQVNGGHLVSIESSGEADFVAHLIAQKIKSAKIHVWIGLRAQNKEKQC
                     SIEWSDGSSISYENWIEEESKKCLGVHKATGFRKWENFYCEQRDPFVCEA"
     sig_peptide     50. .118
                     /gene="acp-a"
BASE COUNT      173 a    148 c    185 g    150 t
ORIGIN

Alignment Scores:
Pred. No.:                4.25e-15        Length:         656
Score:                    155.00          Matches:        26
Percent Similarity:       89.66%          Conservative:   0
Best Local Similarity:    89.66%          Mismatches:     3
Query Match:              91.18%          Indels:         0
DB:                       5               Gaps:           0

US-09-938-114-3 (1-29) x AB036880 (1-656)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||||||||   ||||||||||||||||||||||||||||||   ||||||||||||
Db        119 GATTGTTCCTCTAGTTGGTCCTCCTATGAAGGGCATTGCTATAAGGCCTTCAAACAATCT 178

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   ||||||||||||||||
Db        179 AAGACCTGGGCAGATGCAGAGAGCTTC 205

RESULT 4
AF540647
LOCUS       AF540647                 688 bp    mRNA    linear   VRT 26-SEP-2002
DEFINITION  Deinagkistrodon acutus agglucetin-beta 1 subunit precursor, mRNA,
            complete cds.
ACCESSION   AF540647
VERSION     AF540647.1  GI:23321262
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 688)
  AUTHORS   Wang,W.J. and Huang,T.F.
  TITLE     A novel tetrameric venom protein, agglucetin from Agkistrodon
            acutus, acts as a glycoprotein Ib agonist
  JOURNAL   Thromb. Haemost. 86 (4), 1077-1086 (2001)
  MEDLINE   21542539
   PUBMED   11686327
REFERENCE   2  (bases 1 to 688)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Molecular structure and functional characterization of agglucetin,
            a tetrameric glycoprotein Ib-binding protein, from Formosan pit
            viper
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 688)
  AUTHORS   Wang,W.-J., Ling,Q.-D. and Huang,T.-F.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2002) Department of Nursing, Chang Gung Institute
            of Technology, 261, Wen-Hwa 1 Rd., Kwei-Shan, Tao-Yuan, Taiwan
            333-03, R.O.C.
FEATURES             Location/Qualifiers
     source          1. .688
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
     CDS             61. .501
                     /note="platelet agglutination inducer; glycoprotein Ib
                     agonist"
                     /codon_start=1
                     /product="agglucetin-beta 1 subunit precursor"
                     /protein_id="AAN23126.1"
                     /db_xref="GI:23321263"
                     /translation="MGRVIFVSFGLLVVFLSLSGTAADCPSEWSSYEGHCYKAFKQSK
                     TWADAEKFCTQQHKGSHLASFHSSEEADFVVTLTTPSLKTDLVWIGLKNIWNGCYWKW
                     SDGTKLDYKDWREQFECLVSRTVNNEWLSMDCGTTYSFVCKFQA"
     sig_peptide     61. .129
     mat_peptide     130. .498
                     /product="agglucetin-beta 1 subunit"
BASE COUNT      181 a    173 c    172 g    162 t
ORIGIN

Alignment Scores:
Pred. No.:                5.8e-14         Length:         688
Score:                    148.00          Matches:        24
Percent Similarity:       86.21%          Conservative:   1
Best Local Similarity:    82.76%          Mismatches:     4
Query Match:              87.06%          Indels:         0
DB:                       5               Gaps:           0

US-09-938-114-3 (1-29) x AF540647 (1-688)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||:::||||||||||||||||||||||||||||||   ||||||||||||
Db        130 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGGCCTTCAAACAATCT 189

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        190 AAGACCTGGGCAGATGCAGAGAAATTC 216

RESULT 5
AF176421
LOCUS       AF176421                 458 bp    mRNA    linear   VRT 19-JAN-2000
DEFINITION  Deinagkistrodon acutus agkisacutacin B chain mRNA, complete cds.
ACCESSION   AF176421
VERSION     AF176421.1  GI:6715114
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 458)
  AUTHORS   Cheng,X., Qian,Y., Liu,Q., Li,B.X.Y., Ding,J., Xu,Z., Huang,W. and
```

```
          Liu,J.
  TITLE   Agkisacutacin, a new fibrinolytic & anti-platelet protein from
          Agkistrodon acutus venom
  JOURNAL Unpublished
REFERENCE 2  (bases 1 to 458)
  AUTHORS Cheng,X., Liu,J., Li,B.X.Y. and Qian,Y.
  TITLE   Direct Submission
  JOURNAL Submitted (08-AUG-1999) Dept. of Biochemistry and Molecular
          Biology, University of Science and Technology of China, School of
          Life-Science, Huangshan Road, Hefei, Anhui 230027, China
FEATURES             Location/Qualifiers
     source          1. .458
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /tissue_type="venom"
     CDS             3. .443
                     /note="agkisacutacin consists of two heterologous subunits
                     (A and B chains) linked by an inter-subunit disulfide
                     bond"
                     /codon_start=1
                     /product="agkisacutacin B chain"
                     /protein_id="AAF26287.1"
                     /db_xref="GI:6715115"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCPSDWSSYEGHCYKPFDEPK
                     TWADAEKFCTQQHKGSHLASFHSSEEADFVVTLTTPSLKTDLVWIGLKNIWNGCYWKW
                     SDGTKLDYKDWREQFECLVSRTVNNEWLSMDCGTTCSFVCKFQA"
BASE COUNT      118 a    109 c    116 g    115 t
ORIGIN

Alignment Scores:
Pred. No.:              2.98e-12        Length:         458
Score:                  136.00          Matches:        22
Percent Similarity:     79.31%          Conservative:   1
Best Local Similarity:  75.86%          Mismatches:     6
Query Match:            80.00%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AF176421 (1-458)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||||   |||   :::
Db         72 GATTGTCCCTCTGATTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 131

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 6
AX427201
LOCUS       AX427201                483 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 1 from Patent WO0214364.
ACCESSION   AX427201
VERSION     AX427201.1  GI:21530563
KEYWORDS    .
SOURCE      Sistrurus miliarius (pigmy rattlesnake)
  ORGANISM  Sistrurus miliarius
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Sistrurus.
REFERENCE   1
  AUTHORS   Sheppard,P.O. and Bishop,P.D.
  TITLE     Rattlesnake venom gland proteins
  JOURNAL   Patent: WO 0214364-A 1 21-FEB-2002;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Sistrurus miliarius"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8758"
     CDS             <3. .458
                     /note="Zsnk2"
                     /codon_start=1
                     /protein_id="CAD35498.1"
                     /db_xref="GI:21530564"
                     /translation="IFVSFGLLVVFLSLSGTGADCPSDWSSYDQHCYKVFSELKTWDD
                     AESFCYTQHRDSRLASIHSSEEEAFVGKLASQTLKFTSMWIGLKDLWKECKWQWSDDT
                     KLDYKAWTRRPYCTVMVVKTDRIFWFNRGCEKTVSFVCKFQARSGDPAV"
BASE COUNT      129 a    112 c    123 g    119 t
ORIGIN

Alignment Scores:
Pred. No.:              6.57e-12        Length:         483
Score:                  134.00          Matches:        22
Percent Similarity:     82.76%          Conservative:   2
Best Local Similarity:  75.86%          Mismatches:     5
Query Match:            78.82%          Indels:         0
DB:                     6               Gaps:           0

US-09-938-114-3 (1-29) x AX427201 (1-483)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||:::   ||||||||||||||||||   :::
Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||||||||
Db        120 AAAACCTGGGATGATGCAGAGAGTTTC 146

RESULT 7
AB019616
LOCUS       AB019616                664 bp    mRNA    linear   VRT 14-NOV-1998
DEFINITION  Agkistrodon blomhoffi mRNA for mamushigin beta, complete cds.
ACCESSION   AB019616
VERSION     AB019616.1  GI:3882118
KEYWORDS    mamushigin beta.
SOURCE      Agkistrodon blomhoffi
  ORGANISM  Agkistrodon blomhoffi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Agkistrodon.
REFERENCE   1  (sites)
  AUTHORS   Sakurai,Y., Fujimura,Y., Kokubo,T., Imamura,K., Kawasaki,T.,
            Handa,M., Suzuki,M., Matsui,T., Titani,K. and Yoshioka,A.
  TITLE     The cDNA cloning and molecular characterization of a snake venom
            platelet glycoprotein Ib-binding protein, mamushigin, from
            Agkistrodon halys blomhoffii venom
  JOURNAL   Thromb. Haemost. 79 (6), 1199-1207 (1998)
  MEDLINE   98319530
   PUBMED   9657448
REFERENCE   2  (bases 1 to 664)
  AUTHORS   Sakurai,Y. and Fujimura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-NOV-1998) Yoshihiko Sakurai, Nara Medical University,
            Department of Blood Transfusion Medicine; Shijo-cho 840, Kashihara,
            Nara 634-8522, Japan (E-mail:ysakurai@nmu-gw.cc.naramed-u.ac.jp,
            Tel:81-744-22-3051(ex.3288), Fax:81-744-29-0771)
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Agkistrodon blomhoffi"
                     /mol_type="mRNA"
                     /db_xref="taxon:61300"
                     /note="synonym:Agkistrodon halys blomhoffii"
     CDS             71. .511
                     /codon_start=1
                     /product="mamushigin beta"
                     /protein_id="BAA34425.1"
                     /db_xref="GI:3882119"
                     /translation="MGRFIFLSFGLLVVFVSLSGTGADCPSDWSSYEGHCYRVFQKEM
                     TWEDAEKFCTQQRKESHLVSFHSSEEVDFVVSMTWPILKYDFVWIGLNNIWNECMVEW
                     TDGTRLSHNAWITESECIAAKTTDNQWLSRPCSRTYNVVCKFQE"
BASE COUNT      174 a    153 c    173 g    164 t
ORIGIN

Alignment Scores:
Pred. No.:              9.35e-12        Length:         664
```

```
Score:                  134.00          Matches:        21
Percent Similarity:     82.76%          Conservative:   3
Best Local Similarity:  72.41%          Mismatches:     5
Query Match:            78.82%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AB019616 (1-664)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||::|||||::::::
Db        140 GATTGTCCCTCTGATTGGTCCTCCTATGAAGGGCATTGCTACAGGGTCTTCCAAAAAGAG 199

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                 ||||||   |||||||||   |||
Db        200 ATGACCTGGGAAGATGCAGAGAAATTC 226

RESULT 8
AY091761
LOCUS       AY091761                478 bp    mRNA    linear   VRT 13-MAY-2002
DEFINITION  Deinagkistrodon acutus clone 2101 ACF 1/2 B-chain mRNA, complete
            cds.
ACCESSION   AY091761
VERSION     AY091761.1  GI:20562942
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     B chain of ACF 1/2 from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 478)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2101"
     CDS             1. .441
                     /note="C-type lectin family member"
                     /codon_start=1
                     /product="ACF 1/2 B-chain"
                     /protein_id="AAM22789.1"
                     /db_xref="GI:20562943"
                     /translation="MGRFIFVSFGLLVVFLSLSGTGADCPSDWSSYEGHCYKPFNEPK
                     NWADAENFCTQQHTGGHLVSFQSTEEADFVVKLAFQTFDYGIFWMGLSNIWNQCNWQW
                     SNAAMLKYTDWAEESYCVYFKSTNNKWRSITCRMIANFVCEFQA"
BASE COUNT      127 a    107 c    125 g    119 t
ORIGIN

Alignment Scores:
Pred. No.:              9.36e-12        Length:         478
Score:                  133.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            78.24%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AY091761 (1-478)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||||||||   |||   :::
Db         70 GATTGTCCCTCTGATTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCAATGAACCG 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                 |||   |||   |||||||||:::|||
Db        130 AAAAACTGGGCGGATGCAGAGAATTTC 156

RESULT 9
AB036881
LOCUS       AB036881                592 bp    mRNA    linear   VRT 23-MAY-2002
DEFINITION  Deinagkistrodon acutus acp-b mRNA for anticoagulant protein-B,
            complete cds.
ACCESSION   AB036881
VERSION     AB036881.1  GI:12060180
KEYWORDS    anticoagulant protein-B.
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1
  AUTHORS   Tani,A., Ogawa,T., Nose,T., Nikandrov,N.N., Deshimaru,M.,
            Chijiwa,T., Chang,C.C., Fukumaki,Y. and Ohno,M.
  TITLE     Characterization, primary structure and molecular evolution of
            anticoagulant protein from Agkistrodon actus venom
  JOURNAL   Toxicon 40 (6), 803-813 (2002)
  MEDLINE   21975857
REFERENCE   2  (bases 1 to 592)
  AUTHORS   Ogawa,T. and Tani,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JAN-2000) Tomohisa Ogawa, Tohoku University, Graduate
            School of Agricultural Science; 1-1, Tsutusmidori Amamiyamachi,
            Aoba-ku, Sendai 981-8555, Japan (E-mail:ogawa@biochem.tohoku.ac.jp,
            URL:http://www.agri.tohoku.ac.jp/hoozo/, Tel:81-22-717-8808,
            Fax:81-22-717-8807)
FEATURES             Location/Qualifiers
     source          1. .592
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /tissue_type="venom gland"
     gene            1. .592
                     /gene="acp-b"
     CDS             5. .445
                     /gene="acp-b"
                     /codon_start=1
                     /product="anticoagulant protein-B"
                     /protein_id="BAB20441.1"
                     /db_xref="GI:12060181"
                     /translation="MGRFIFVSFGLLVLFLSLSGTAADCPSDWSSYEGHCYKPFNEPK
                     NWADAENFCTQQHTGSHLVSFQSTEEADFVVKLAFQTFDYGIFWMGLSKIWNQCNWQW
                     SNAAMLKYTDWAEESYCVYFKSTNNKWRSITCRMIANFVCEFQA"
     sig_peptide     5. .73
                     /gene="acp-b"
BASE COUNT      149 a    146 c    149 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              1.19e-11        Length:         592
Score:                  133.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            78.24%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AB036881 (1-592)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||||||   |||   :::
Db         74 GATTGTCCCTCTGATTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCAATGAACCG 133

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                 |||   |||   |||||||||:::|||
Db        134 AAAAACTGGGCGGATGCAGAGAATTTC 160

RESULT 10
TFLFIXB
LOCUS       TFLFIXB                 698 bp    mRNA    linear   VRT 06-FEB-1999
```

```
DEFINITION  Trimeresurus flavoviridis mRNA for factor IX/factor X binding
            protein B chain, complete cds.
ACCESSION   D83332
VERSION     D83332.1  GI:1402641
KEYWORDS    factor IX/factor X binding protein B chain.
SOURCE      Trimeresurus flavoviridis
  ORGANISM  Trimeresurus flavoviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Trimeresurus.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   Matsuzaki,R., Yoshiara,E., Yamada,M., Shima,K., Atoda,H. and
            Morita,T.
  TITLE     cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant
            protein from snake venom
  JOURNAL   Biochem. Biophys. Res. Commun. 220 (2), 382-387 (1996)
  MEDLINE   96184662
   PUBMED   8645314
REFERENCE   2  (bases 1 to 698)
  AUTHORS   Morita,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1996) Takashi Morita, Meiji College of Pharmacy,
            Biochemistry; 1-22-1 Yatocho, Tanashi, Tokyo 188, Japan
            (Tel:0424-21-0101(ex.429), Fax:0424-21-1489)
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="Trimeresurus flavoviridis"
                     /mol_type="mRNA"
                     /db_xref="taxon:88087"
     5'UTR           1. .73
     CDS             74. .514
                     /codon_start=1
                     /product="factor IX/factor X binding protein B chain"
                     /protein_id="BAA11888.1"
                     /db_xref="GI:1402642"
                     /translation="MGRFIFMSFGFLVVFLSLSGTAADCPSDWSSYEGHCYKPFSEPK
                     NWADAENFCTQQHAGGHLVSFQSSEEADFVVKLAFQTFGHSIFWMGLSNVWNQCNWQW
                     SNAAMLRYKAWAEESYCVYFKSTNNKWRSRACRMMAQFVCEFQA"
     3'UTR           515. .698
     polyA_signal    679. .684
BASE COUNT      167 a    185 c    177 g    169 t
ORIGIN

Alignment Scores:
Pred. No.:              1.42e-11        Length:         698
Score:                  133.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            78.24%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x TFLFIXB (1-698)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||||   |||   :::
Db        143 GATTGTCCCTCTGATTGGTCCTCTTATGAAGGGCATTGCTACAAGCCCTTCAGTGAACCC 202

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||   |||   |||||||||:::|||
Db        203 AAAAACTGGGCCGATGCAGAGAATTTC 229

RESULT 11
AY293866
LOCUS       AY293866                369 bp    DNA     linear   VRT 09-JUN-2003
DEFINITION  Deinagkistrodon acutus agkisasin-b gene, partial cds.
ACCESSION   AY293866
VERSION     AY293866.1  GI:31559054
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 369)
  AUTHORS   Zha,X.D., Ren,B., Liu,J. and Xu,K.S.
  TITLE     Genomic DNA sequence of b chain of Agkisasin, a C-type lectin-like
            protein from Agkistrodon acutus venom, and its evolutionary
            significance
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 369)
  AUTHORS   Zha,X.D., Ren,B., Liu,J. and Xu,K.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAY-2003) Biology, School of Life Science, Anhui
            University, 3 Feixi Road, Hefei 230039, China
FEATURES             Location/Qualifiers
     source          1. .369
                     /organism="Deinagkistrodon acutus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:36307"
     mRNA            <1. .>369
                     /product="agkisasin-b"
     CDS             <1. .>369
                     /note="C-type lectin-like protein"
                     /codon_start=1
                     /product="agkisasin-b"
                     /protein_id="AAP50528.1"
                     /db_xref="GI:31559055"
                     /translation="DCPSDWSSYEGHCYKPFNELKNWADAENFCTQQHTGSHLVSFQS
                     TEEADFVVKLAFQTFDYGIFWMGLSNIWNQCNWQWSNAAMLKYTDWAEESYCVYFKST
                     NNKWRSITCRMIANFVCEFQA"
BASE COUNT      107 a     83 c     88 g     91 t
ORIGIN

Alignment Scores:
Pred. No.:              1.01e-11        Length:         369
Score:                  132.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            77.65%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AY293866 (1-369)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||||||||||||||||   |||   :::
Db          1 GATTGTCCTTCTGATTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCAATGAACTG 60

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||   |||   |||||||||:::|||
Db         61 AAAAACTGGGCGGATGCAGAGAATTTC 87

RESULT 12
AR259041
LOCUS       AR259041                454 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 12 from patent US 6489451.
ACCESSION   AR259041
VERSION     AR259041.1  GI:27309526
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Li,B.X. and Cheng,X.
  TITLE     Antithrombosis enzyme from the snake venom of agkistrodon acutus
  JOURNAL   Patent: US 6489451-A 12 03-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="unknown"
BASE COUNT      114 a    107 c    114 g    109 t     10 others
ORIGIN

Alignment Scores:
Pred. No.:              1.27e-11        Length:         454
Score:                  132.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
```

```
Query Match:            77.65%      Indels:        0
DB:                     6           Gaps:          0

US-09-938-114-3 (1-29) x AR259041 (1-454)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||:::|||||||||||||||||||||||||||||   |||   :::
Db         72 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 131

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 13
AY091756
LOCUS       AY091756                478 bp    mRNA    linear   VRT 13-MAY-2002
DEFINITION  Deinagkistrodon acutus clone 2100488 agkisacutacin B-chain mRNA,
            complete cds.
ACCESSION   AY091756
VERSION     AY091756.1  GI:20562934
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     B chain of agkisacutacin from Deinagkistrodon acutus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 478)
  AUTHORS   Yu,H., Xiang,K., Wang,Y. and Liu,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Dept. of Biochemistry and Molecular
            Biology, University of Science and Technology of China, School of
            Life Science, Huangshan Road, Hefei, Anhui 230027, P.R. China
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /clone="2100488"
     CDS             1. .441
                     /note="C-type lectin family member"
                     /codon_start=1
                     /product="agkisacutacin B-chain"
                     /protein_id="AAM22785.1"
                     /db_xref="GI:20562935"
                     /translation="MGRFIFVSFGLLVVFLSLSGTAADCPSEWSSYEGHCYKPFDEPK
                     TWADAEKFCTQQHKGSHLASFHSSEEADFVVTLTTPSLKTDLVWIGLKNIWNGCYWKW
                     SDGTKLDYKDWREQFECLVSRTVNNEWLSMDCGTTCSFVCKFQA"
BASE COUNT      126 a    109 c    126 g    117 t
ORIGIN

Alignment Scores:
Pred. No.:              1.35e-11        Length:         478
Score:                  132.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            77.65%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AY091756 (1-478)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||:::|||||||||||||||||||||||||||||   |||   :::
Db         70 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        130 AAGACCTGGGCAGATGCAGAGAAATTC 156

RESULT 14
AF350324
LOCUS       AF350324                574 bp    mRNA    linear   VRT 25-MAR-2001
DEFINITION  Deinagkistrodon acutus agkisasin-b mRNA, complete cds.
ACCESSION   AF350324
VERSION     AF350324.1  GI:13445903
KEYWORDS    .
SOURCE      Deinagkistrodon acutus
  ORGANISM  Deinagkistrodon acutus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Deinagkistrodon.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Zha,X. and Xu,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2001) Biochemical and Genetic Engeneering Drugs,
            National Institute for Control of Phamaceuticals and Biological
            Products, Tiantan Ximen, Xili Street 2, Beijing 10050, China
FEATURES             Location/Qualifiers
     source          1. .574
                     /organism="Deinagkistrodon acutus"
                     /mol_type="mRNA"
                     /db_xref="taxon:36307"
                     /tissue_type="venom gland"
                     /lab_host="Escherichia coli JM109"
     CDS             1. .375
                     /note="IX/X-BP family member; coagulant and anticoagulant"
                     /codon_start=1
                     /product="agkisasin-b"
                     /protein_id="AAK26430.1"
                     /db_xref="GI:13445904"
                     /translation="MDCPSDWSSYEGHCYKPFNELKNWADAENFCTQQHTGSHLVSFQ
                     STEEADFVVKLAFQTFDYGIFWMGLSNIWNQCNWQWSNAAMLKYTDWAEESYCVYFKS
                     TNNKWRSITCRMIANFVCEFQA"
BASE COUNT      170 a    134 c    133 g    137 t
ORIGIN

Alignment Scores:
Pred. No.:              1.65e-11        Length:         574
Score:                  132.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            77.65%          Indels:         0
DB:                     5               Gaps:           0

US-09-938-114-3 (1-29) x AF350324 (1-574)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||||||||||||||||||||||||||||||||||   |||   :::
Db          4 GATTGTCCTTCTGATTGGTCCTCTTATGAAGGGCATTGCTACAAGCCCTTCAATGAACTG 63

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||   |||   |||||||||:::|||
Db         64 AAAAACTGGGCGGATGCAGAGAATTTC 90

RESULT 15
AF197915
LOCUS       AF197915                634 bp    mRNA    linear   VRT 01-NOV-2000
DEFINITION  Gloydius halys halyxin B-chain precursor, mRNA, complete cds.
ACCESSION   AF197915
VERSION     AF197915.1  GI:11066255
KEYWORDS    .
SOURCE      Gloydius halys (halys viper)
  ORGANISM  Gloydius halys
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
            Viperidae; Crotalinae; Gloydius.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Koo,B.H., Sohn,Y.D., Kim,D.S., Jang,Y.S. and Chung,K.H.
  TITLE     A Novel Coagulation Factor Xa Inhibitor from Korean Snake
            (Agkistrodon halys) venom
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 634)
  AUTHORS   Sohn,Y.D., Koo,B.H., Kim,D.S., Jang,Y.S. and Chung,K.H.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1999) Cardiovascular center, Yonsei University
            College of Med., 134 Shinchon-dong, Seoul 120-752, Republic of
            Korea
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Gloydius halys"
                     /mol_type="mRNA"
                     /db_xref="taxon:8714"
                     /tissue_type="venom"
     CDS             24. .464
                     /note="coagulation factor Xa inhibitor"
                     /codon_start=1
                     /product="halyxin B-chain precursor"
                     /protein_id="AAG28522.1"
                     /db_xref="GI:11066256"
                     /translation="MGRFIFLSFGLLVVFLSLSGTGADCPSGWSSYEGHCYKPFNEQK
                     NWADAENFCTQQHTGGHLVSFHSTEEADFVVKLAFQNFGHGIFWMGLSNVWNQCSWQW
                     SSAAKLKYEAWAEESYCVYFKSTNNKWRSRACRMEAYFVCEFQA"
     sig_peptide     24. .92
     mat_peptide     93. .461
                     /product="halyxin B-chain"
BASE COUNT      180 a    142 c    157 g    155 t
ORIGIN

Alignment Scores:
Pred. No.:                1.15e-10        Length:         634
Score:                    127.00          Matches:        20
Percent Similarity:       75.86%          Conservative:   2
Best Local Similarity:    68.97%          Mismatches:     7
Query Match:              74.71%          Indels:         0
DB:                       5               Gaps:           0

US-09-938-114-3 (1-29) x AF197915 (1-634)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||||||||||||||||||   |||   :::
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAACCATTTAATGAACAA 152

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||   |||   |||||||||:::|||
Db        153 AAAAATTGGGCCGATGCAGAGAATTTC 179


Search completed: December  8, 2003, 16:08:45
Job time : 1231.93 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 12:39:59 ; Search time 99.5943 Seconds
(without alignments)
786.026 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091000_22839/app_query.fasta_1.725
-DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_0_@runat_08122003_091000_22839 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	78.8	483	24	AAD32053	Pigmy rattlesnake
2	132	77.6	454	25	ABX93674	cDNA encoding Dein
3	127	74.7	633	22	AAI71877	Snake venom blood
4	127	74.7	721	24	AAD32055	Pigmy rattlesnake
5	119	70.0	453	24	AAD32054	Pigmy rattlesnake
6	119	70.0	583	24	AAL42016	Korean adder snake
7	119	70.0	601	22	AAI71876	Snake venom blood
8	115	67.6	456	24	AAD32056	Pigmy rattlesnake
9	115	67.6	544	25	ABX93668	cDNA encoding Dein
10	111	65.3	690	16	AAQ89309	Snake venom antith
11	111	65.3	690	21	AAC61144	DNA encoding a sna
12	111	65.3	690	24	AAK99834	DNA encoding the a
13	110	64.7	580	24	AAD32057	Pigmy rattlesnake
14	103	60.6	432	24	AAD32058	Pigmy rattlesnake
15	99	58.2	632	24	AAL42015	Korean adder snake
16	97	57.1	725	24	AAD32059	Pigmy rattlesnake
17	94	55.3	716	17	AAT64829	Snake venom protei
18	91	53.5	474	24	AAD32060	Pigmy rattlesnake
19	84	49.4	492	22	AAI64154	Human Zlec1 encodi
20	84	49.4	492	24	AAD27975	Human lectin Zlec1
21	84	49.4	5191	15	AAQ57710	Neurocan DNA. Hom
22	82	48.2	495	24	ABK32999	DNA encoding novel
23	82	48.2	629	22	AAI64153	Human Zlec1 encodi
24	82	48.2	629	24	AAD27974	Human lectin Zlec1
25	82	48.2	655	22	AAC84316	Human EXCS encodin
26	82	48.2	2652	25	ABZ81727	Rat mutant brain-e
27	82	48.2	2652	25	ABZ81728	Rat brain-enriched
28	82	48.2	3153	22	AAF29464	Murine brevidin cD
29	80	47.1	790	11	AAQ05300	Bovine pancreatic
30	79	46.5	3259	16	AAT05627	Bovine brevican co
c 31	76	44.7	634	21	AAC69541	Human secreted pro
32	76	44.7	2878	25	AAD48133	Human chondroitin
33	76	44.7	3476	22	AAA91017	Human secreted pro
34	76	44.7	3476	24	ABK69986	cDNA encoding huma
35	75	44.1	2087	24	ABA02881	Human versican iso
36	75	44.1	2465	24	ABK62593	Rat sequence diffe
37	75	44.1	4302	21	AAC77694	Human cancer assoc
38	75	44.1	4932	25	ABX63334	Human cDNA #334 di
39	75	44.1	8224	12	AAQ12261	Versican gene. Ho
40	75	44.1	8224	24	ABT11088	Human breast cance
41	75	44.1	8224	25	ACC50121	Breast cancer asso
42	75	44.1	11185	24	ABN96814	Gene #3312 used to
43	75	44.1	11185	24	ABL62702	Colon adenocarcino
44	75	44.1	12319	24	AAS94985	Human DNA sequence
45	74	43.5	857	22	AAS31230	Human cDNA encodin

ALIGNMENTS

RESULT 1
AAD32053
ID AAD32053 standard; DNA; 483 BP.
XX
AC AAD32053;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein, Zsnk2 gene.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; gene; Zsnk2; ds.
XX
OS Sistrurus miliarius.
XX
FH Key Location/Qualifiers
FT CDS 3..458

```
FT                   /*tag=  a
FT                   /product= "Zsnk2 protein"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   sig_peptide     3..59
FT                   /*tag=  b
FT   mat_peptide     60..455
FT                   /*tag=  c
FT                   /product= "Mature Zsnk2 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20178.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 71-72; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk2 gene.
XX
SQ   Sequence 483 BP; 129 A; 112 C; 123 G; 119 T; 0 other;

Alignment Scores:
Pred. No.:                1.47e-10        Length:          483
Score:                    134.00          Matches:         22
Percent Similarity:       82.76%          Conservative:    2
Best Local Similarity:    75.86%          Mismatches:      5
Query Match:              78.82%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-3 (1-29) x AAD32053 (1-483)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||:::   ||||||||||||||||||   :::
Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   ||||||||||||||
Db        120 AAAACCTGGGATGATGCAGAGAGTTTC 146
```

```
RESULT 2
ABX93674
ID   ABX93674 standard; cDNA; 454 BP.
XX
AC   ABX93674;
XX
DT   02-JUN-2003  (first entry)
XX
DE   cDNA encoding Deinagkistrodon acutus antithrombosis enzyme beta chain.
XX
KW   Antithrombosis; ss; PCR; beta chain; fibrin hydrolysis; blood clot;
KW   platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW   myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW   angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW   ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW   unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW   peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW   thrombosis.
XX
OS   Deinagkistrodon acutus.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..443
FT                   /*tag=  a
FT                   /product= "Antithrombosis enzyme beta chain"
FT                   /transl_except= (pos:216..218,aa:Xaa)
FT                   /transl_except= (pos:219..221,aa:Xaa)
FT                   /transl_except= (pos:222..224,aa:Xaa)
FT                   /transl_except= (pos:225..227,aa:Xaa)
FT                   /note= "Xaa = unknown"
FT   sig_peptide     3..71
FT                   /*tag=  b
FT                   /note= "Leader peptide"
FT   mat_peptide     72..440
FT                   /*tag=  c
FT                   /note= "Mature beta chain"
FT   misc_feature    217..226
FT                   /*tag=  d
FT                   /note= "Unsequenced region could be 10-20 nucleotides in
FT                   size"
XX
PN   US6489451-B1.
XX
PD   03-DEC-2002.
XX
PF   10-APR-1998;   98US-0058740.
XX
PR   10-APR-1997;   97US-043886P.
XX
PA   (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI   Li BX,  Cheng X;
XX
DR   WPI; 2003-352116/33.
DR   P-PSDB; ABU08799.
XX
PT   New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT   preventing and treating vaso-occulusive and thromboembolic disorders,
PT   including myocardial infarction, restenosis, cerebral thrombosis and
PT   unstable angina  -
XX
PS   Disclosure; Fig 1; 19pp; English.
XX
CC   The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC   enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC   prevents platelet aggregation. The anti-thrombosis enzyme was
CC   administered to rabbits intravenously. Thrombosis was determined before
CC   and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC   Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC   following administration at 0.005 micro/kg and this activity was
CC   increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC   composition is useful for preventing and treating vaso-occulusive and
CC   thromboembolic disorders, including myocardial infarction, restenosis,
```

```
CC   angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC   diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC   thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC   peripheral arterial occulsion, stroke. It is also useful for treating
CC   atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC   diseases. The present sequence represents cDNA encoding the
CC   Deinagkistrodon acutus antithrombosis enzyme beta chain.
XX
SQ   Sequence 454 BP; 114 A; 107 C; 114 G; 109 T; 10 other;

Alignment Scores:
Pred. No.:                2.67e-10      Length:          454
Score:                    132.00        Matches:         21
Percent Similarity:       79.31%        Conservative:    2
Best Local Similarity:    72.41%        Mismatches:      6
Query Match:              77.65%        Indels:          0
DB:                       25            Gaps:            0

US-09-938-114-3 (1-29) x ABX93674 (1-454)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||    |||:::|||||||||||||||||||||||||||||    |||    :::
Db         72 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 131

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 3
AAI71877
ID   AAI71877 standard; DNA; 633 BP.
XX
AC   AAI71877;
XX
DT   10-JAN-2002  (first entry)
XX
DE   Snake venom blood anticoagulant halyxin B chain coding sequence.
XX
KW   Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW   thrombosis; ds.
XX
OS   Agkistrodon halys.
XX
PN   KR2001049671-A.
XX
PD   15-JUN-2001.
XX
PF   29-JUN-2000; 2000KR-0036591.
XX
PR   29-JUN-1999;   99KR-0025105.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Jang YS,  Jung GH,  Kim DS,  Koo BH,  Son YD;
XX
DR   WPI; 2001-637330/73.
DR   P-PSDB; AAM51544.
XX
PT   Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS   Claim 1; Page 11; 21pp; Korean.
XX
CC   The invention relates to halyxin, a novel protein with very strong
CC   blood anticoagulation activity. The protein was separated from snake
CC   venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC   be used in the treatment of thrombogenesis. The present sequence
CC   encodes the B chain of halyxin.
XX
SQ   Sequence 633 BP; 179 A; 142 C; 157 G; 155 T; 0 other;

Alignment Scores:
Pred. No.:                2.19e-09      Length:          633
Score:                    127.00        Matches:         20
Percent Similarity:       75.86%        Conservative:    2
Best Local Similarity:    68.97%        Mismatches:      7
Query Match:              74.71%        Indels:          0
DB:                       22            Gaps:            0

US-09-938-114-3 (1-29) x AAI71877 (1-633)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||    |||    ||||||||||||||||||||||||||||    |||    :::
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAAACCATTTAATGAACAA 152

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||    |||    |||||||||:::|||
Db        153 AAAAATTGGGCCGATGCAGAGAATTTC 179

RESULT 4
AAD32055
ID   AAD32055 standard; DNA; 721 BP.
XX
AC   AAD32055;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk3 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk3; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             91..549
FT                   /*tag=  a
FT                   /product= "Zsnk3 protein"
FT   sig_peptide     91..159
FT                   /*tag=  b
FT   mat_peptide     160..546
FT                   /*tag=  c
FT                   /product= "Mature Zsnk3 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20179.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 73-74; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
```

CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein, Zsnk3 gene.
XX
SQ Sequence 721 BP; 192 A; 167 C; 203 G; 159 T; 0 other;

Alignment Scores:

Pred. No.:	2.56e-09	Length:	721
Score:	127.00	Matches:	21
Percent Similarity:	79.31%	Conservative:	2
Best Local Similarity:	72.41%	Mismatches:	6
Query Match:	74.71%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-3 (1-29) x AAD32055 (1-721)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||:::  ||||||||||:::||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        220 AAGACGTGGGACGATGCAGAGAGGTTC 246
```

RESULT 5
AAD32054
ID AAD32054 standard; DNA; 453 BP.
XX
AC AAD32054;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk2, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 72-73; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet
CC aggregation system. The polypeptides, which affect blood coagulation and
CC platelet aggregation system, are useful in therapy and diagnostics. The
CC polypeptides are also useful as an educational tool in laboratory
CC practical kits for courses related to genetics and molecular biology,
CC protein chemistry and antibody production and analysis. The
CC polynucleotide or polypeptide can be used as standards or as unknowns
CC for testing purposes. The polypeptides are also useful in identifying
CC proteins by western blotting, protein purification, determining the
CC weight of expressed polypeptides as a ratio to total protein expressed,
CC identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC tags, mass spectrometry, circular dichroism to determine conformation or
CC affinity chromatography columns to purify the protein, cloning or
CC sequencing. The present sequence is Sistrurus miliarius venom gland
CC protein Zsnk2, degenerate nucleic acid.
XX
SQ Sequence 453 BP; 74 A; 32 C; 84 G; 70 T; 193 other;

Alignment Scores:

Pred. No.:	2.17e-08	Length:	453
Score:	119.00	Matches:	18
Percent Similarity:	68.97%	Conservative:	2
Best Local Similarity:	62.07%	Mismatches:	9
Query Match:	70.00%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-3 (1-29) x AAD32054 (1-453)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||      |||:::  ||||||||||||||||||  :::
Db         58 GAYTGYCCNWSNGAYTGGWSNWSNTAYGAYCARCAYTGYTAYAARGTNTTYWSNGARYTN 117

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        118 AARACNTGGGAYGAYGCNGARWSNTTY 144
```

RESULT 6
AAL42016
ID AAL42016 standard; cDNA; 583 BP.
XX
AC AAL42016;
XX
DT 16-MAY-2002 (first entry)
XX
DE Korean adder snake venom salmorin B chain protein cDNA sequence.
XX
KW Korean adder; ss; salmorin protein B chain; snake venom;
KW fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW thrombin binding; blood coagulation.
XX
OS Agkistrodon halys brevicaudus.
XX

FH	Key	Location/Qualifiers
FT	CDS	7..444
FT		/*tag= a
FT		/product= "Salmorin B chain protein"
FT	sig_peptide	7..75
FT		/*tag= b
FT		/note= "Signal peptide"
FT	mat_peptide	76..441
FT		/*tag= c
FT		/note= "Mature Salmorin B chain protein"
FT	3'UTR	445..570
FT		/*tag= d
FT	polyA_signal	571..583
FT		/*tag= e

XX
PN WO200214514-A1.
XX
PD 21-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-KR01277.
XX
PR 27-JUL-2000; 2000KR-0043470.
XX

PA (BIOB-) BIOBUD CO LTD.
XX
PI Chung K, Kim D, Koh Y;
XX
DR WPI; 2002-241907/29.
DR P-PSDB; AAO14521.
XX
PT New salmorin protein derived from venom of Korean adder Agkistrodon
PT halys brevicaudus, useful for treating thrombosis by repressing
PT fibrinogen clotting through repression of activation of prothrombin
PT into thrombin -
XX
PS Claim 1; Fig 1B; 30pp; English.
XX
CC The invention comprises the nucleotide and protein sequences of a
CC salmorin protein derived from the venom of Korean adder. Salmorin protein
CC is composed of an A chain and a B chain, and has inhibitory activity
CC against fibrinogen clotting. The salmorin protein of the invention is
CC useful for treating thrombosis, as it represses fibrinogen clotting
CC potently by binding to prothrombin and thrombin so as to delay blood
CC coagulation. The present cDNA sequence encodes the Korean adder salmorin
CC B chain protein.
XX
SQ Sequence 583 BP; 151 A; 137 C; 150 G; 145 T; 0 other;

Alignment Scores:

Pred. No.:	2.96e-08	Length:	583
Score:	119.00	Matches:	19
Percent Similarity:	68.97%	Conservative:	1
Best Local Similarity:	65.52%	Mismatches:	9
Query Match:	70.00%	Indels:	0
DB:	24	Gaps:	0

US-09-938-114-3 (1-29) x AAL42016 (1-583)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||||||||||||||||:::|||   |||
Db         76 GATTGTCCCTCTGGTTGGTCCTCTTATGAAGGGCATTGCTACAAGCTCTTCAATCAACAG 135

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                 |||      ||||||   |||
Db        136 TCTCAGTGGGCCCATGCAGAGAAATTC 162
```

RESULT 7
AAI71876
ID AAI71876 standard; DNA; 601 BP.
XX
AC AAI71876;
XX
DT 10-JAN-2002 (first entry)
XX
DE Snake venom blood anticoagulant halyxin A chain coding sequence.
XX
KW Snake; pit viper; venom; halyxin; anticoagulant; thrombogenesis;
KW thrombosis; ds.
XX
OS Agkistrodon halys.
XX
PN KR2001049671-A.
XX
PD 15-JUN-2001.
XX
PF 29-JUN-2000; 2000KR-0036591.
XX
PR 29-JUN-1999; 99KR-0025105.
XX
PA (BIOB-) BIOBUD CO LTD.
XX
PI Jang YS, Jung GH, Kim DS, Koo BH, Son YD;
XX
DR WPI; 2001-637330/73.
DR P-PSDB; AAM51543.
XX
PT Halyxin as blood anticoagulation protein separated from snake venom -
XX
PS Claim 1; Page 9; 21pp; Korean.
XX
CC The invention relates to halyxin, a novel protein with very strong
CC blood anticoagulation activity. The protein was separated from snake
CC venom of Agkistrodon halys brevicaudus (a Korean pit viper) and can
CC be used in the treatment of thrombogenesis. The present sequence
CC encodes the A chain of halyxin.
XX
SQ Sequence 601 BP; 166 A; 133 C; 163 G; 139 T; 0 other;

Alignment Scores:

Pred. No.:	3.08e-08	Length:	601
Score:	119.00	Matches:	19
Percent Similarity:	72.41%	Conservative:	2
Best Local Similarity:	65.52%	Mismatches:	8
Query Match:	70.00%	Indels:	0
DB:	22	Gaps:	0

US-09-938-114-3 (1-29) x AAI71876 (1-601)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||||||||||||||||:::|||
Db         93 GATTGTCCCTCTGGTTGGTCCTCCTATGAAGGGCATTGCTACAACATCTTCCATCTATTC 152

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              ||||||||   :::||||||   |||
Db        153 AAGACCTGGGCAGAAGCAGAGAGGTTC 179
```

RESULT 8
AAD32056
ID AAD32056 standard; DNA; 456 BP.
XX
AC AAD32056;
XX
DT 18-JUN-2002 (first entry)
XX
DE Pigmy rattlesnake venom gland protein Zsnk3, degenerate nucleic acid.
XX
KW Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW platelet aggregation; ds.
XX
OS Sistrurus miliarius.
XX
PN WO200214364-A2.
XX
PD 21-FEB-2002.
XX
PF 13-AUG-2001; 2001WO-US25310.
XX
PR 14-AUG-2000; 2000US-225072P.
PR 14-AUG-2000; 2000US-225087P.
PR 15-AUG-2000; 2000US-225489P.
PR 15-AUG-2000; 2000US-225490P.
PR 20-DEC-2000; 2000US-356997P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sheppard PO, Bishop PD;
XX
DR WPI; 2002-269180/31.
DR P-PSDB; AAE20179.
XX
PT New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT affect blood coagulation and platelet aggregation system, useful in
PT therapy and diagnostics, or as tools in the study of genetics or
PT molecular biology -
XX
PS Disclosure; Page 74; 79pp; English.
XX
CC The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC venom gland proteins, which affect blood coagulation and platelet

```
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein Zsnk3, degenerate nucleic acid.
XX
SQ   Sequence 456 BP; 72 A; 31 C; 92 G; 66 T; 195 other;

Alignment Scores:
Pred. No.:              8.49e-08        Length:         456
Score:                  115.00          Matches:        18
Percent Similarity:     65.52%          Conservative:   1
Best Local Similarity:  62.07%          Mismatches:     10
Query Match:            67.65%          Indels:         0
DB:                     24              Gaps:           0

US-09-938-114-3 (1-29) x AAD32056 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||        |||       |||:::   |||||||||   ||||||||||||
Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        130 AARACNTGGGAYGAYGCNGARMGNTTY 156

RESULT 9
ABX93668
ID   ABX93668 standard; cDNA; 544 BP.
XX
AC   ABX93668;
XX
DT   02-JUN-2003  (first entry)
XX
DE   cDNA encoding Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
KW   Antithrombosis; ss; PCR; alpha chain; fibrin hydrolysis; blood clot;
KW   platelet aggregation; vaso-occulusive disorder; thromboembolic disorder;
KW   myocardial infarction; restenosis; cancer; neurodegenerative disease;
KW   angiopathic thrombosis; cerebral thrombosis; thromboangiitis obliterans;
KW   ischaemic cerebral vascular disease; unstable angina; acute thrombosis;
KW   unstable stenocardia; pulmonary embolism; deep vein thrombosis; oedema;
KW   peripheral arterial occulsion; stroke; atherosclerosis; inflammation;
KW   thrombosis.
XX
OS   Deinagkistrodon acutus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..363
FT                   /*tag=  a
FT                   /product= "Antithrombosis enzyme alpha chain"
FT                   /partial
FT                   /transl_except= (pos:244..246,aa:Lys)
FT                   /note= "No start codon given"
XX
PN   US6489451-B1.
XX
PD   03-DEC-2002.
XX
PF   10-APR-1998;   98US-0058740.
XX
PR   10-APR-1997;   97US-043886P.
XX
PA   (HEFE-) HEFEI SIU FUNG USTC PHARM CO LTD.
XX
PI   Li BX,  Cheng X;
XX
DR   WPI; 2003-352116/33.
DR   P-PSDB; ABU08798.
XX
PT   New purified Agkistrodon actus anti-thrombosis enzyme, useful for
PT   preventing and treating vaso-occulusive and thromboembolic disorders,
PT   including myocardial infarction, restenosis, cerebral thrombosis and
PT   unstable angina  -
XX
PS   Example 3; Column 25-26; 19pp; English.
XX
CC   The invention relates to a new Deinagkistrodon actus anti-thrombosis
CC   enzyme, where the enzyme hydrolyses fibrin, dissolves blood clots and
CC   prevents platelet aggregation. The anti-thrombosis enzyme was
CC   administered to rabbits intravenously. Thrombosis was determined before
CC   and after the administration at 0.5, 1.0, 2.0 and 3.0 hours by the
CC   Chandler method. The enzyme showed anti-thrombosis activity at 0.5 hour
CC   following administration at 0.005 micro/kg and this activity was
CC   increased significantly at 1.0 hour and at 0.01 micro /kg. The
CC   composition is useful for preventing and treating vaso-occulusive and
CC   thromboembolic disorders, including myocardial infarction, restenosis,
CC   angiopathic thrombosis, cerebral thrombosis, ischaemic cerebral vascular
CC   diseases, unstable angina, acute thrombosis, unstable stenocardia,
CC   thromboangiitis obliterans, pulmonary embolism, deep vein thrombosis,
CC   peripheral arterial occulsion, stroke. It is also useful for treating
CC   atherosclerosis, oedema and inflammation, cancer and neurodegenerative
CC   diseases. The present sequence represents cDNA encoding the
CC   Deinagkistrodon acutus antithrombosis enzyme alpha chain.
XX
SQ   Sequence 544 BP; 158 A; 122 C; 143 G; 121 T; 0 other;

Alignment Scores:
Pred. No.:              1.05e-07        Length:         544
Score:                  115.00          Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            67.65%          Indels:         0
DB:                     25              Gaps:           0

US-09-938-114-3 (1-29) x ABX93668 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

RESULT 10
AAQ89309
ID   AAQ89309 standard; cDNA; 690 BP.
XX
AC   AAQ89309;
XX
DT   25-MAR-2003  (updated)
DT   28-NOV-1995  (first entry)
XX
DE   Snake venom antithrombotic oligopeptide cDNA.
XX
KW   Antithrombotic peptide; snake venom; platelet binding inhibition;
KW   von Willebrand factors; Crotalus horridus horridus; ds.
XX
OS   Crotalus horridus horridus.
XX
FH   Key             Location/Qualifiers
FT   CDS             66..515
FT                   /*tag=  a
XX
PN   WO9508573-A1.
XX
PD   30-MAR-1995.
XX
PF   21-SEP-1994;   94WO-JP01555.
XX
```

PR 22-SEP-1993; 93JP-0236975.
XX
PA (AJIN) AJINOMOTO KK.
XX
PI Fukuchi N, Ishii K, Kito M, Kobayashi T, Nagano M;
PI Tanaka A, Yamamoto H, Yoshimoto R;
XX
DR WPI; 1995-139559/18.
DR P-PSDB; AAR71981.
XX
PT Single-chain antithrombotic peptide - obtained by cleaving an
PT oligopeptide from snake venom to break inter-chain di:sulphide
PT bonds but preserve intra-chain di:sulphide bonds
XX
PS Example 2; Pages 47-48; 84pp; Japanese.
XX
CC AAR71978 and AAR71979 are snake venom derived antithrombotic peptides,
CC specifically from the snake venom oligopeptide AAR71981, encoded by
CC AAQ89309. These peptides have the advantage of avoiding significant
CC thrombocytopenia when administered at the minimun dose, for in vivo
CC inhibition of platelet von Willebrand factor binding.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:

Pred. No.:	5.47e-07	Length:	690
Score:	111.00	Matches:	18
Percent Similarity:	72.41%	Conservative:	3
Best Local Similarity:	62.07%	Mismatches:	8
Query Match:	65.29%	Indels:	0
DB:	16	Gaps:	0

US-09-938-114-3 (1-29) x AAQ89309 (1-690)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   ||||||||||||:::   :::||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTC 227
```

RESULT 11
AAC61144
ID AAC61144 standard; DNA; 690 BP.
XX
AC AAC61144;
XX
DT 07-FEB-2001 (first entry)
XX
DE DNA encoding a snake venom derived protein.
XX
KW Subunit peptide production; snake venom; rattlesnake; thrombolytic;
KW von Willebrand's factor; blood platelet-inhibitory activity; ds.
XX
OS Crotalus horridus horridus.
XX
PN WO200059926-A1.
XX
PD 12-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-JP02127.
XX
PR 02-APR-1999; 99JP-0096073.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Fukuchi N, Kageyama S, Kito M, Kayahara T, Yamamoto H;
XX
DR WPI; 2000-664985/64.
DR P-PSDB; AAY85628.
XX
PT Producing physiologically-active subunit peptides originating in
PT polymer proteins by denaturation and specific separation, with lower
PT antigenicity but improved solubility and stability, e.g. blood
PT platelet-binding inhibitors -
XX
PS Disclosure; Page 46; 51pp; Japanese.
XX
CC This invention relates to a method for the production of a subunit
CC peptide originating from a polymer protein with disulphide bonds within
CC and between subunits. The method comprises denaturing the protein or its
CC subunit using a protein denaturing agent in a solution, removing the
CC agent in the presence of a polyoxyalkyl polyether which reacts with a
CC thiol group and unwinds the subunit, and separating the polyoxyalkyl
CC polyether-bound subunit peptide. The method can be used for producing
CC physiologically-active subunit peptides for polymer proteins e.g. snake
CC venom-originated dimer peptide with blood platelet-inhibitory activity
CC on von Willebrand's factor. The peptides produced have platelet-binding
CC inhibitory, and thrombolytic activity. The present sequence represents
CC DNA encoding rattlesnake protein used in an example illustrating the
CC method of the invention.
XX
SQ Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:

Pred. No.:	5.47e-07	Length:	690
Score:	111.00	Matches:	18
Percent Similarity:	72.41%	Conservative:	3
Best Local Similarity:	62.07%	Mismatches:	8
Query Match:	65.29%	Indels:	0
DB:	21	Gaps:	0

US-09-938-114-3 (1-29) x AAC61144 (1-690)

```
Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   ||||||||||||:::   :::||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTC 227
```

RESULT 12
AAK99834
ID AAK99834 standard; DNA; 690 BP.
XX
AC AAK99834;
XX
DT 19-JUL-2002 (first entry)
XX
DE DNA encoding the antithrombotic wild-type rattlesnake protein.
XX
KW Thrombolytic; glycoprotein Ib-binding protein; snake venom; rattlesnake;
KW antithrombotic; glycoprotein Ib; long half life; low antigenicity;
KW drug efficacy; gene; ds.
XX
OS Crotalus horridus horridus.
XX
FH Key Location/Qualifiers
FT CDS 66..515
FT /*tag= a
FT /product= "Antithrombotic wild-type rattlesnake protein"
XX
PN EP1195384-A1.
XX
PD 10-APR-2002.
XX
PF 04-OCT-2001; 2001EP-0123277.
XX
PR 04-OCT-2000; 2000JP-0305279.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Fukuchi N, Kito M, Kayahara T, Futaki F, Ishikawa K, Suzuki E;

```
PI   Gondoh K,  Shimba N,  Yamada N;
XX
DR   WPI; 2002-364482/40.
DR   P-PSDB; AAO20974.
XX
PT   Glycoprotein Ib-binding protein, useful for treating thrombosis,
PT   comprises specific mutations in protein originating from snake venom  -
XX
PS   Example 1; Page 25; 49pp; English.
XX
CC   The invention relates to a glycoprotein Ib-binding protein, originating
CC   from snake venom, comprising specific mutations and antithrombotic
CC   activity. Glycoprotein Ib-binding protein is used in a drug having
CC   antithrombotic activity. Glycoprotein Ib-binding protein has a high
CC   binding activity to glycoprotein Ib, a long half life/drug efficacy
CC   retention in blood, and low antigenicity. This polynucleotide sequence
CC   represents DNA of the wild-type rattlesnake protein of the invention.
XX
SQ   Sequence 690 BP; 167 A; 180 C; 170 G; 173 T; 0 other;

Alignment Scores:
Pred. No.:                5.47e-07       Length:          690
Score:                    111.00         Matches:         18
Percent Similarity:       72.41%         Conservative:    3
Best Local Similarity:    62.07%         Mismatches:      8
Query Match:              65.29%         Indels:          0
DB:                       24             Gaps:            0

US-09-938-114-3 (1-29) x AAK99834 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||    |||   |||||||||||:::   :::|||||||||    |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTC 227

RESULT 13
AAD32057
ID   AAD32057 standard; DNA; 580 BP.
XX
AC   AAD32057;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein, Zsnk4 gene.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; gene; Zsnk4; ds.
XX
OS   Sistrurus miliarius.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..437
FT                   /*tag=  a
FT                   /product= "Zsnk4 protein"
FT                   /note= "CDS does not include start codon"
FT                   /partial
FT   sig_peptide     3..29
FT                   /*tag=  b
FT   mat_peptide     30..434
FT                   /*tag=  c
FT                   /product= "Mature Zsnk4 protein"
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
DR   P-PSDB; AAE20180.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Claim 5; Page 75; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein, Zsnk4 gene.
XX
SQ   Sequence 580 BP; 154 A; 133 C; 163 G; 130 T; 0 other;

Alignment Scores:
Pred. No.:                6.2e-07        Length:          580
Score:                    110.00         Matches:         17
Percent Similarity:       75.86%         Conservative:    5
Best Local Similarity:    58.62%         Mismatches:      7
Query Match:              64.71%         Indels:          0
DB:                       24             Gaps:            0

US-09-938-114-3 (1-29) x AAD32057 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||||||||   :::|||:::   :::||||||:::|||   ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              :::||||||   |||||||||   |||
Db         96 AGGACGTGGGACGATGCAGAGAGGTTC 122

RESULT 14
AAD32058
ID   AAD32058 standard; DNA; 432 BP.
XX
AC   AAD32058;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Pigmy rattlesnake venom gland protein Zsnk4, degenerate nucleic acid.
XX
KW   Pigmy rattlesnake; venom gland protein; blood coagulation; therapy;
KW   platelet aggregation; ds.
XX
OS   Sistrurus miliarius.
XX
PN   WO200214364-A2.
XX
PD   21-FEB-2002.
```

```
XX
PF   13-AUG-2001; 2001WO-US25310.
XX
PR   14-AUG-2000; 2000US-225072P.
PR   14-AUG-2000; 2000US-225087P.
PR   15-AUG-2000; 2000US-225489P.
PR   15-AUG-2000; 2000US-225490P.
PR   20-DEC-2000; 2000US-356997P.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Sheppard PO,  Bishop PD;
XX
DR   WPI; 2002-269180/31.
XX
PT   New pigmy rattlesnake (Sistrurus miliarius) venom gland proteins, which
PT   affect blood coagulation and platelet aggregation system, useful in
PT   therapy and diagnostics, or as tools in the study of genetics or
PT   molecular biology  -
XX
PS   Disclosure; Page 76; 79pp; English.
XX
CC   The invention relates to new pigmy rattlesnake (Sistrurus miliarius)
CC   venom gland proteins, which affect blood coagulation and platelet
CC   aggregation system. The polypeptides, which affect blood coagulation and
CC   platelet aggregation system, are useful in therapy and diagnostics. The
CC   polypeptides are also useful as an educational tool in laboratory
CC   practical kits for courses related to genetics and molecular biology,
CC   protein chemistry and antibody production and analysis. The
CC   polynucleotide or polypeptide can be used as standards or as unknowns
CC   for testing purposes. The polypeptides are also useful in identifying
CC   proteins by western blotting, protein purification, determining the
CC   weight of expressed polypeptides as a ratio to total protein expressed,
CC   identifying peptide cleavage sites, coupling amino and carboxyl terminal
CC   tags, mass spectrometry, circular dichroism to determine conformation or
CC   affinity chromatography columns to purify the protein, cloning or
CC   sequencing. The present sequence is Sistrurus miliarius venom gland
CC   protein Zsnk4, degenerate nucleic acid.
XX
SQ   Sequence 432 BP; 91 A; 34 C; 84 G; 55 T; 168 other;

Alignment Scores:
Pred. No.:                4.62e-06        Length:          432
Score:                    103.00          Matches:         16
Percent Similarity:       65.52%          Conservative:    3
Best Local Similarity:    55.17%          Mismatches:      10
Query Match:              60.59%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-3 (1-29) x AAD32058 (1-432)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                 ||||||   |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTY 120

RESULT 15
AAL42015
ID   AAL42015 standard; cDNA; 632 BP.
XX
AC   AAL42015;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Korean adder snake venom salmorin A chain protein cDNA sequence.
XX
KW   Korean adder; ss; salmorin protein A chain; snake venom;
KW   fibrinogen clotting inhibition; thrombosis; prothrombin binding;
KW   thrombin binding; blood coagulation.
XX
OS   Agkistrodon halys brevicaudus.
XX
FH   Key             Location/Qualifiers
FT   CDS             34..498
FT                   /*tag=  a
FT                   /product= "Salmorin A chain protein"
FT   sig_peptide     34..102
FT                   /*tag=  b
FT                   /note= "Signal peptide"
FT   mat_peptide     103..495
FT                   /*tag=  c
FT                   /note= "Mature Salmorin A chain protein"
FT   3'UTR           499..622
FT                   /*tag=  d
FT   polyA_signal    623..632
FT                   /*tag=  e
XX
PN   WO200214514-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JUL-2001; 2001WO-KR01277.
XX
PR   27-JUL-2000; 2000KR-0043470.
XX
PA   (BIOB-) BIOBUD CO LTD.
XX
PI   Chung K,  Kim D,  Koh Y;
XX
DR   WPI; 2002-241907/29.
DR   P-PSDB; AAO14520.
XX
PT   New salmorin protein derived from venom of Korean adder Agkistrodon
PT   halys brevicaudus, useful for treating thrombosis by repressing
PT   fibrinogen clotting through repression of activation of prothrombin
PT   into thrombin  -
XX
PS   Claim 1; Fig 1A; 30pp; English.
XX
CC   The invention comprises the nucleotide and protein sequences of a
CC   salmorin protein derived from the venom of Korean adder. Salmorin protein
CC   is composed of an A chain and a B chain, and has inhibitory activity
CC   against fibrinogen clotting. The salmorin protein of the invention is
CC   useful for treating thrombosis, as it represses fibrinogen clotting
CC   potently by binding to prothrombin and thrombin so as to delay blood
CC   coagulation. The present cDNA sequence encodes the Korean adder salmorin
CC   A chain protein.
XX
SQ   Sequence 632 BP; 162 A; 140 C; 182 G; 148 T; 0 other;

Alignment Scores:
Pred. No.:                2.86e-05        Length:          632
Score:                    99.00           Matches:         16
Percent Similarity:       60.71%          Conservative:    1
Best Local Similarity:    57.14%          Mismatches:      11
Query Match:              58.24%          Indels:          0
DB:                       24              Gaps:            0

US-09-938-114-3 (1-29) x AAL42015 (1-632)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||   |||   |||   |||       ||||||||||:::   |||   |||
Db        112 TGTCCCTCTGGTTGGGGCAGCAATAATGGTCATTGCTACCAGGCCTTCAATCAACGGATG 171

Qy         22 ThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        172 ACCTGGGAAGATGCAGAGAGGTTC 195


Search completed: December  8, 2003, 14:04:47
Job time : 100.594 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 13:49:04 ; Search time 22.8686 Seconds
(without alignments)
559.725 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091001_22878/app_query.fasta_1.725
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09938114_@CGN_1_1_95_@runat_08122003_091001_22878 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	132	77.6	454	4	US-09-058-740-12	Sequence 12, Appl
2	115	67.6	544	4	US-09-058-740-1	Sequence 1, Appli
3	111	65.3	690	2	US-08-612-840A-7	Sequence 7, Appli
4	84	49.4	5191	1	US-08-340-428B-1	Sequence 1, Appli
5	84	49.4	5191	5	PCT-US93-07306-1	Sequence 1, Appli
6	79	46.5	3259	5	PCT-US95-03747-1	Sequence 1, Appli
7	75	44.1	8224	6	5180808-1	Patent No. 5180808
8	73	42.9	402	3	US-08-543-246B-10	Sequence 10, Appl
9	73	42.9	648	3	US-08-543-246B-14	Sequence 14, Appl
10	73	42.9	1740	2	US-09-055-095-2	Sequence 2, Appli
11	73	42.9	1755	3	US-08-543-246B-8	Sequence 8, Appli
12	73	42.9	1841	4	US-09-996-243-318	Sequence 318, App
13	73	42.9	4588	3	US-08-840-062-1	Sequence 1, Appli
14	73	42.9	4771	3	US-08-840-062-3	Sequence 3, Appli
15	72	42.4	4978	1	US-08-220-603A-1	Sequence 1, Appli
16	71	41.8	522	1	US-07-778-156-4	Sequence 4, Appli
17	71	41.8	522	2	US-08-422-166-4	Sequence 4, Appli
18	71	41.8	600	5	PCT-US93-10418-1	Sequence 1, Appli
19	71	41.8	602	3	US-09-385-982-257	Sequence 257, App
20	71	41.8	797	2	US-08-464-637-1	Sequence 1, Appli
21	71	41.8	797	2	US-08-822-261-5	Sequence 5, Appli
22	71	41.8	797	4	US-09-226-852-5	Sequence 5, Appli
23	71	41.8	798	1	US-07-778-156-12	Sequence 12, Appl
24	71	41.8	798	2	US-08-822-261-6	Sequence 6, Appli
25	71	41.8	798	2	US-08-422-166-12	Sequence 12, Appl
26	71	41.8	798	3	US-09-146-969-2	Sequence 2, Appli
27	71	41.8	798	4	US-09-226-852-6	Sequence 6, Appli
28	69	40.6	1104	3	US-09-111-470-1	Sequence 1, Appli
29	69	40.6	2076	4	US-09-489-847-51	Sequence 51, Appl
30	68	40.0	614	2	US-08-729-103-2	Sequence 2, Appli
31	68	40.0	1114	2	US-08-468-413-1	Sequence 1, Appli
32	68	40.0	1114	3	US-09-162-508-1	Sequence 1, Appli
33	68	40.0	1114	5	PCT-US95-07169-1	Sequence 1, Appli
34	67	39.4	43804	4	US-09-171-461-1	Sequence 1, Appli
35	66	38.8	360	3	US-08-543-246B-4	Sequence 4, Appli
36	66	38.8	405	3	US-08-543-246B-3	Sequence 3, Appli
37	66	38.8	405	3	US-08-543-246B-7	Sequence 7, Appli
38	66	38.8	645	3	US-08-543-246B-12	Sequence 12, Appl
39	66	38.8	693	3	US-08-543-246B-13	Sequence 13, Appl
40	66	38.8	699	3	US-08-543-246B-11	Sequence 11, Appl
41	66	38.8	1222	3	US-08-543-246B-5	Sequence 5, Appli
42	66	38.8	1223	4	US-09-016-434-1347	Sequence 1347, Ap
43	66	38.8	1333	3	US-08-543-246B-15	Sequence 15, Appl
44	66	38.8	1387	3	US-08-543-246B-1	Sequence 1, Appli
45	66	38.8	1643	4	US-09-517-605-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-058-740-12
; Sequence 12, Application US/09058740
; Patent No. 6489451
;  GENERAL INFORMATION:
;       APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                  Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                  Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                  Jin-Guo Ding, Fang Rong, Yan Liu and
;                  Hui-Ran Chen
;       TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                           VENOM OF AGKISTRODON ACUTUS
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 633 West Fifth Street
;                    Suite 4700
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: U.S.A.
;            ZIP: 90071-2066
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                         storage
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: IBM P.C. DOS 5.0
;            SOFTWARE: FastSEQ for Windows 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/058,740
;            FILING DATE: 10-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: <Unknown>
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Chen, Anthony C.
```

```
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 12:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 454 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          FEATURE:
;             NAME/KEY:          Coding Sequence
;             LOCATION:          3...440
;             OTHER INFORMATION:    "N" stands for any base.
;             "Xaa" stands for any amino acid.
;          SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-058-740-12

Alignment Scores:
Pred. No.:                1.73e-11        Length:         454
Score:                    132.00          Matches:        21
Percent Similarity:       79.31%          Conservative:   2
Best Local Similarity:    72.41%          Mismatches:     6
Query Match:              77.65%          Indels:         0
DB:                       4               Gaps:           0

US-09-938-114-3 (1-29) x US-09-058-740-12 (1-454)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||:::|||||||||||||||||||||||||||||   |||   :::
Db         72 GATTGTCCCTCTGAGTGGTCCTCCTATGAAGGGCATTGCTACAAGCCCTTCGATGAACCT 131

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        132 AAGACCTGGGCAGATGCAGAGAAATTC 158

RESULT 2
US-09-058-740-1
; Sequence 1, Application US/09058740
; Patent No. 6489451
;   GENERAL INFORMATION:
;          APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                     Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                     Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                     Jin-Guo Ding, Fang Rong, Yan Liu and
;                     Hui-Ran Chen
;          TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                              VENOM OF AGKISTRODON ACUTUS
;          NUMBER OF SEQUENCES: 12
;          CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;          COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/058,740
;             FILING DATE: 10-Apr-1998
;             CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: <Unknown>
;             FILING DATE: <Unknown>
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Chen, Anthony C.
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;    INFORMATION FOR SEQ ID NO: 1:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 544 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-058-740-1

Alignment Scores:
Pred. No.:                8.2e-09         Length:         544
Score:                    115.00          Matches:        20
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              67.65%          Indels:         0
DB:                       4               Gaps:           0

US-09-938-114-3 (1-29) x US-09-058-740-1 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

RESULT 3
US-08-612-840A-7
; Sequence 7, Application US/08612840A
; Patent No. 5856126
;  GENERAL INFORMATION:
;    APPLICANT:  FUKUCHI, Naoyuki
;    APPLICANT:  YAMAMOTO, Hiroshi
;    APPLICANT:  NAGANO, Mitsuyo
;    APPLICANT:  KITO, Morikazu
;    APPLICANT:  TANAKA, Akiko
;    APPLICANT:  ISHII, Koichi
;    APPLICANT:  KOBAYASHI, Tsuyoshi
;    APPLICANT:  YOSHIMOTO, Ryota
;    TITLE OF INVENTION:  PEPTIDE HAVING ANTI-THROMBUS ACTIVITY AND
;    TITLE OF INVENTION:  METHOD OF PRODUCING THE SAME
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Oblon, Spivak, McClelland, Maier & Neustadt, P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,840A
;      FILING DATE:  20-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-236975
;      FILING DATE:  22-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5856126man F. Oblon
;      REGISTRATION NUMBER:  24,618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  690 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Crotalus horridus horridus
;       STRAIN:
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  66..512
US-08-612-840A-7

Alignment Scores:
Pred. No.:                4.44e-08        Length:         690
Score:                    111.00          Matches:        18
Percent Similarity:       72.41%          Conservative:   3
Best Local Similarity:    62.07%          Mismatches:     8
Query Match:              65.29%          Indels:         0
DB:                       2               Gaps:           0

US-09-938-114-3 (1-29) x US-08-612-840A-7 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||||||||||:::   :::||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTC 227

RESULT 4
US-08-340-428B-1
; Sequence 1, Application US/08340428B
; Patent No. 5648465
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,428B
;      FILING DATE:  14 No. 5648465ember 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/922,911
;      FILING DATE:  03 August 1992
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  Margolis=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  5191 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  77..3847
US-08-340-428B-1

Alignment Scores:
Pred. No.:                0.00661         Length:         5191
Score:                    84.00           Matches:        11
Percent Similarity:       57.69%          Conservative:   4
Best Local Similarity:    42.31%          Mismatches:     11
Query Match:              49.41%          Indels:         0
DB:                       1               Gaps:           0

US-09-938-114-3 (1-29) x US-08-340-428B-1 (1-5191)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||         |||      ::::::||||||||||||:::   |||         :::
Db       3161 TGCGACCGTGGCTGGCACAAATTCCAGGGCCACTGCTACCGCTACTTTGCTCATCGGCGG 3220

Qy         22 ThrTrpThrAspAlaGlu 27
                |||   |||||||||
Db       3221 GCCTGGGAGGACGCAGAG 3238

RESULT 5
PCT-US93-07306-1
; Sequence 1, Application PC/TUS9307306
;  GENERAL INFORMATION:
;    APPLICANT:  MARGOLIS, Richard U.
;    APPLICANT:  RAUCH, Uwe
;    APPLICANT:  MARGOLIS, Renee K.
;    TITLE OF INVENTION:  CLONING, EXPRESSION AND USES FOR A
;    TITLE OF INVENTION:  NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/07306
;      FILING DATE:  03-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/922,911
;      FILING DATE:  03-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Townsend, Guy K.
;      REGISTRATION NUMBER:  34,033
;      REFERENCE/DOCKET NUMBER:  Margolis=1A PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  5191 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  77..3847
```

```
PCT-US93-07306-1

Alignment Scores:
Pred. No.:                0.00661         Length:          5191
Score:                    84.00           Matches:         11
Percent Similarity:       57.69%          Conservative:    4
Best Local Similarity:    42.31%          Mismatches:      11
Query Match:              49.41%          Indels:          0
DB:                       5               Gaps:            0

US-09-938-114-3 (1-29) x PCT-US93-07306-1 (1-5191)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||          |||       ::::::|||||||||||||:::   |||       :::
Db       3161 TGCGACCGTGGCTGGCACAAATTCCAGGGCCACTGCTACCGCTACTTTGCTCATCGGCGG 3220

Qy         22 ThrTrpThrAspAlaGlu 27
                |||   |||||||||
Db       3221 GCCTGGGAGGACGCAGAG 3238

RESULT 6
PCT-US95-03747-1
; Sequence 1, Application PC/TUS9503747
;  GENERAL INFORMATION:
;    APPLICANT:  LA JOLLA CANCER RESEARCH FOUNDATION
;    TITLE OF INVENTION:  Brevican, A Glial Cell Proteoglycan
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03747
;      FILING DATE:  27-MAR-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Imbra, Richard J.
;      REGISTRATION NUMBER:  37,643
;      REFERENCE/DOCKET NUMBER:  FP-LJ 1453
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3259 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  112..2848
PCT-US95-03747-1

Alignment Scores:
Pred. No.:                0.0214          Length:          3259
Score:                    79.00           Matches:         11
Percent Similarity:       66.67%          Conservative:    7
Best Local Similarity:    40.74%          Mismatches:      9
Query Match:              46.47%          Indels:          0
DB:                       5               Gaps:            0

US-09-938-114-3 (1-29) x PCT-US95-03747-1 (1-3259)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              ||||||       |||   :::::::::|||   |||||||||   |||         :::
Db       2176 TGCAGCCCCGGTTGGGACGCCTTCCAGGGTGCCTGCTACAAGCACTTTTCTGCCCGAAGG 2235

Qy         22 ThrTrpThrAspAlaGluSer 28
                 :::|||   :::||||||:::
Db       2236 AGCTGGGAGGAGGCGGAGAAC 2256

RESULT 7
5180808-1
;Patent No. 5180808
;    APPLICANT: RUOSLAHTI, ERKKI I.
;    TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID
;SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN
;ANTIBODIES, AND METHODS OF DETECTING THE SAME
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/441,179
;      FILING DATE: 27-NOV-1989
;SEQ ID NO:1:
;      LENGTH: 8224
5180808-1

Alignment Scores:
Pred. No.:                0.27            Length:          8224
Score:                    75.00           Matches:         11
Percent Similarity:       53.85%          Conservative:    3
Best Local Similarity:    42.31%          Mismatches:      12
Query Match:              44.12%          Indels:          0
DB:                       6               Gaps:            0

US-09-938-114-3 (1-29) x 5180808-1 (1-8224)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||          |||       ::::::|||   |||||||||   |||         :::
Db       6810 TGTGACTATGGCTGGCACAAATTCCAAGGGCAGTGCTACAAATACTTTGCCCATCGACGC 6869

Qy         22 ThrTrpThrAspAlaGlu 27
                ||||||       ||||||
Db       6870 ACATGGGATGCAGCTGAA 6887

RESULT 8
US-08-543-246B-10
; Sequence 10, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  402 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-543-246B-10

Alignment Scores:
Pred. No.:                0.0132          Length:         402
Score:                    73.00           Matches:        9
Percent Similarity:       66.67%          Conservative:   9
Best Local Similarity:    33.33%          Mismatches:     9
Query Match:              42.94%          Indels:         0
DB:                       3               Gaps:           0

US-09-938-114-3 (1-29) x US-08-543-246B-10 (1-402)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||       :::|||       |||:::   :::||||||:::   |||   :::||||||
Db         49 TGTCCTAAAAACTGGATATGTTACAAAAATAACTGCTACCAATTTTTTGATGAGAGTAAA 108

Qy         22 ThrTrpThrAspAlaGluSer 28
                |||   ::::::::::::
Db        109 AACTGGTATGAGAGCCAGGCT 129

RESULT 9
US-08-543-246B-14
; Sequence 14, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  648 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-543-246B-14

Alignment Scores:
Pred. No.:                0.0238          Length:         648
Score:                    73.00           Matches:        9
Percent Similarity:       66.67%          Conservative:   9
Best Local Similarity:    33.33%          Mismatches:     9
Query Match:              42.94%          Indels:         0
DB:                       3               Gaps:           0

US-09-938-114-3 (1-29) x US-08-543-246B-14 (1-648)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||       :::|||       |||:::   :::||||||:::   |||   :::||||||
Db        295 TGTCCTAAAAACTGGATATGTTACAAAAATAACTGCTACCAATTTTTTGATGAGAGTAAA 354

Qy         22 ThrTrpThrAspAlaGluSer 28
                |||   ::::::::::::
Db        355 AACTGGTATGAGAGCCAGGCT 375

RESULT 10
US-09-055-095-2
; Sequence 2, Application US/09055095
; Patent No. 5945308
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Y. Tom
;    APPLICANT:  Patterson, Chandra
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Sather, Susan
;    TITLE OF INVENTION:  HUMAN OXIDIZED LDL RECEPTOR
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/055,095
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0500 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1740 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  LUNGNOT09
;      CLONE:  1355922
US-09-055-095-2

Alignment Scores:
Pred. No.:                0.0803          Length:          1740
Score:                    73.00           Matches:         10
Percent Similarity:       60.71%          Conservative:    7
Best Local Similarity:    35.71%          Mismatches:      11
Query Match:              42.94%          Indels:          0
DB:                       2               Gaps:            0

US-09-938-114-3 (1-29) x US-09-055-095-2 (1-1740)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||:::     |||      :::      :::||||||:::   :::|||   ||||||
Db        446 TGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAA 505

Qy         22 ThrTrpThrAspAlaGluSerPhe 29
              :::|||   |||   :::   |||
Db        506 AGTTGGGAGGACTGTAAATATTTC 529

RESULT 11
US-08-543-246B-8
; Sequence 8, Application US/08543246B
; Patent No. 6262244
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA and amino acid sequence specific for
;    TITLE OF INVENTION:  natural killer cells
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Michael W. Glynn
;      ADDRESSEE:  No. 6262244artis Corporation
;      STREET:  564 Morris Avenue
;      CITY:  Summit,
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07901-1027
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/543,246B
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/676,663
;      FILING DATE:  28-MAR-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/02469
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/122,514
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kassenoff, Melvyn M.
;      REGISTRATION NUMBER:  26,389
;      REFERENCE/DOCKET NUMBER:  118-7704/PCT/CONT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-522-6927
;      TELEFAX:  908-522-6955
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1755 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  339..986
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  339..986
US-08-543-246B-8

Alignment Scores:
Pred. No.:                0.0811          Length:          1755
Score:                    73.00           Matches:         9
Percent Similarity:       66.67%          Conservative:    9
Best Local Similarity:    33.33%          Mismatches:      9
Query Match:              42.94%          Indels:          0
DB:                       3               Gaps:            0

US-09-938-114-3 (1-29) x US-08-543-246B-8 (1-1755)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||      :::|||      |||:::   :::||||||:::   |||   :::||||||
Db        633 TGTCCTAAAAACTGGATATGTTACAAAAATAACTGCTACCAATTTTTTGATGAGAGTAAA 692

Qy         22 ThrTrpThrAspAlaGluSer 28
              |||   ::::::::::::
Db        693 AACTGGTATGAGAGCCAGGCT 713

RESULT 12
US-09-996-243-318
; Sequence 318, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi,Avi J.
;  APPLICANT:  Baker,Kevin P.
;  APPLICANT:  Botstein,David
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Eaton,Dan L.
;  APPLICANT:  Ferrara,Napoleone
;  APPLICANT:  Fong,Sherman
;  APPLICANT:  Gerber,Hanspeter
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,J.Christopher
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Kljavin,Ivar J.
;  APPLICANT:  Napier,Mary A.
;  APPLICANT:  Pan,James
;  APPLICANT:  Paoni,Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Watanabe, Colin K.
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2730P1C13
;  CURRENT APPLICATION NUMBER: US/09/996,243
;  CURRENT FILING DATE:  2001-11-14
;  PRIOR APPLICATION NUMBER: 60/049787
;  PRIOR FILING DATE: 1997-06-16
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/065186
;  PRIOR FILING DATE: 1997-11-12
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066770
;  PRIOR FILING DATE: 1997-11-24
```

; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01

```
;  PRIOR APPLICATION NUMBER: 60/091519
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091626
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091633
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/091978
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/091982
;  PRIOR FILING DATE: 1998-07-07
;  PRIOR APPLICATION NUMBER: 60/092182
;  PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.:                0.086          Length:          1841
Score:                    73.00          Matches:         10
Percent Similarity:       60.71%         Conservative:    7
Best Local Similarity:    35.71%         Mismatches:      11
Query Match:              42.94%         Indels:          0
DB:                       4              Gaps:            0

US-09-938-114-3 (1-29) x US-09-996-243-318 (1-1841)

Qy          2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
              |||:::      |||      :::      :::|||||:::   :::|||   ||||||
Db        552 TGTACAGAACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAA 611

Qy         22 ThrTrpThrAspAlaGluSerPhe 29
              :::|||   |||   :::   |||
Db        612 AGTTGGGAGGACTGTAAATATTTC 635

RESULT 13
US-08-840-062-1
; Sequence 1, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4588 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-840-062-1

Alignment Scores:
Pred. No.:                0.264          Length:          4588
Score:                    73.00          Matches:         7
Percent Similarity:       66.67%         Conservative:    11
Best Local Similarity:    25.93%         Mismatches:      9
Query Match:              42.94%         Indels:          0
DB:                       3              Gaps:            0

US-09-938-114-3 (1-29) x US-08-840-062-1 (1-4588)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          |||      ::::::||||||||||||:::::::      :::
Db       1242 GAATGTGACCCCAGCTGGCAGCCCTTCCAGGGCCACTGCTACCGCCTGCAGGCCGAGAAG 1301

Qy         21 LysThrTrpThrAspAlaGlu 27
              :::::|||   :::::::::
Db       1302 CGCAGCTGGCAGGAGTCCAAG 1322

RESULT 14
US-08-840-062-3
; Sequence 3, Application US/08840062
; Patent No. 6117977
;  GENERAL INFORMATION:
;    APPLICANT:  LASKY, LAURENCE A.
;    APPLICANT:  WU, KAI
;    TITLE OF INVENTION:  TYPE C LECTINS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,062
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Ginger R.
;      REGISTRATION NUMBER:  33,055
;      REFERENCE/DOCKET NUMBER:  P1019R1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-3216
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4771 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-840-062-3

Alignment Scores:
Pred. No.:                0.277          Length:          4771
Score:                    73.00          Matches:         7
Percent Similarity:       66.67%         Conservative:    11
Best Local Similarity:    25.93%         Mismatches:      9
Query Match:              42.94%         Indels:          0
DB:                       3              Gaps:            0

US-09-938-114-3 (1-29) x US-08-840-062-3 (1-4771)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||          |||      ::::::||||||||||||:::::::      :::
Db       1153 GAGTGCGAGCCGAGCTGGCAGCCCTTCCAGGGCCACTGCTACCGCCTGCAGGCCGAGAAG 1212

Qy         21 LysThrTrpThrAspAlaGlu 27
```

```
              :::::::|||   :::::::::
Db       1213 CGCAGCTGGCAGGAGTCCAAG 1233

RESULT 15
US-08-220-603A-1
; Sequence 1, Application US/08220603A
; Patent No. 5612190
;  GENERAL INFORMATION:
;    APPLICANT:  H. Arita et al.
;    TITLE OF INVENTION:  A Bovine Group I Phospholipase A2 Receptor
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  c/o FISH & NEAVE
;      STREET:  1251 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10020
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette. 3.50 inch. 1.4 MB format
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS ver. 2.11
;      SOFTWARE:  Word Perfect (ASCII file)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/220,603A
;      FILING DATE:  30-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 5-72323
;      FILING DATE:  30-March-1993
;      APPLICATION NUMBER:  JP 5-206686
;      FILING DATE:  20-August-1993
;      APPLICATION NUMBER:  JP 6-40177
;      FILING DATE:  10-March-1994
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4978
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  bovine
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  280 to 4668
;      IDENTIFICATION METHOD:  by similarity with known sequence or to
;      IDENTIFICATION METHOD:  an established consensus
US-08-220-603A-1

Alignment Scores:
Pred. No.:                0.414          Length:         4978
Score:                    72.00          Matches:        10
Percent Similarity:       64.00%         Conservative:   6
Best Local Similarity:    40.00%         Mismatches:     9
Query Match:              42.35%         Indels:         0
DB:                       1              Gaps:           0

US-09-938-114-3 (1-29) x US-08-220-603A-1 (1-4978)

Qy           2 CysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSerLys 21
               |||          |||:::   :::      :::|||||||||:::   |||:::   |||
Db        1411 TGTGAGCCTGGCTGGAATCCCCACAACCGTAATTGCTATAAACTGCAGAAAGAAAAAAAG 1470

Qy          22 ThrTrpThrAspAla 26
               ||||||   :::|||
Db        1471 ACCTGGAATGAGGCT 1485


Search completed: December  8, 2003, 17:27:03
Job time : 25.8686 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 8, 2003, 16:08:50 ; Search time 115.337 Seconds
(without alignments)
835.675 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09938114/runat_08122003_091002_22963/app_query.fasta_1.725
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09938114_@CGN_1_1_504_@runat_08122003_091002_22963
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	134	78.8	483	10	US-09-929-230-1	Sequence 1, Appli
2	134	78.8	483	13	US-10-226-420-1	Sequence 1, Appli
3	127	74.7	721	10	US-09-929-230-4	Sequence 4, Appli
4	127	74.7	721	13	US-10-226-420-4	Sequence 4, Appli
5	119	70.0	453	10	US-09-929-230-3	Sequence 3, Appli
6	119	70.0	453	13	US-10-226-420-3	Sequence 3, Appli
7	115	67.6	456	10	US-09-929-230-6	Sequence 6, Appli
8	115	67.6	456	13	US-10-226-420-6	Sequence 6, Appli
9	115	67.6	544	11	US-09-938-114-1	Sequence 1, Appli
10	111	65.3	690	10	US-09-969-763-2	Sequence 2, Appli
11	110	64.7	580	10	US-09-929-230-7	Sequence 7, Appli
12	110	64.7	580	13	US-10-226-420-7	Sequence 7, Appli
13	103	60.6	432	10	US-09-929-230-9	Sequence 9, Appli
14	103	60.6	432	13	US-10-226-420-9	Sequence 9, Appli
15	97	57.1	725	10	US-09-929-230-10	Sequence 10, Appl
16	97	57.1	725	13	US-10-226-420-10	Sequence 10, Appl
17	91	53.5	474	10	US-09-929-230-12	Sequence 12, Appl
18	91	53.5	474	13	US-10-226-420-12	Sequence 12, Appl
19	84	49.4	195	13	US-10-029-386-19055	Sequence 19055, A
20	84	49.4	492	9	US-09-801-438-3	Sequence 3, Appli
21	84	49.4	512	13	US-10-029-386-5299	Sequence 5299, Ap
22	84	49.4	6310	13	US-10-241-220-26	Sequence 26, Appl
23	82	48.2	495	10	US-09-893-737-19	Sequence 19, Appl
24	82	48.2	629	9	US-09-801-438-1	Sequence 1, Appli
25	82	48.2	655	10	US-09-965-528-50	Sequence 50, Appl
26	82	48.2	2652	15	US-10-195-970-4	Sequence 4, Appli
27	82	48.2	2652	15	US-10-195-970-5	Sequence 5, Appli
28	82	48.2	3153	11	US-09-759-130B-343	Sequence 343, App
29	82	48.2	3153	13	US-10-188-495-73	Sequence 73, Appl
30	82	48.2	3153	15	US-10-189-123-73	Sequence 73, Appl
c 31	76	44.7	634	15	US-10-050-704-83	Sequence 83, Appl
c 32	76	44.7	2694	15	US-10-050-704-40	Sequence 40, Appl
33	76	44.7	3476	13	US-10-237-496-51	Sequence 51, Appl
34	76	44.7	3476	13	US-10-242-074-51	Sequence 51, Appl
35	76	44.7	3476	13	US-10-242-505-51	Sequence 51, Appl
36	76	44.7	3476	13	US-10-242-574-51	Sequence 51, Appl
37	76	44.7	3476	13	US-10-243-261-51	Sequence 51, Appl
38	76	44.7	3476	13	US-10-243-282-51	Sequence 51, Appl
39	76	44.7	3476	13	US-10-243-402-51	Sequence 51, Appl
40	76	44.7	3476	13	US-10-243-431-51	Sequence 51, Appl
41	76	44.7	3476	13	US-10-245-164-51	Sequence 51, Appl
42	76	44.7	3476	13	US-10-244-972-51	Sequence 51, Appl
43	76	44.7	3476	13	US-10-197-942-51	Sequence 51, Appl
44	76	44.7	3476	13	US-10-238-196-51	Sequence 51, Appl
45	76	44.7	3476	13	US-10-245-013-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-09-929-230-1
; Sequence 1, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
; FILE REFERENCE: 00-72
; CURRENT APPLICATION NUMBER: US/09/929,230
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Sistrurus miliarius
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(455)
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Zsnk2
US-09-929-230-1

```
Alignment Scores:
Pred. No.:              1.24e-12        Length:         483
Score:                  134.00          Matches:        22
Percent Similarity:     82.76%          Conservative:   2
Best Local Similarity:  75.86%          Mismatches:     5
Query Match:            78.82%          Indels:         0
DB:                     10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-929-230-1 (1-483)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||:::   ||||||||||||||||||   :::
Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   ||||||||||||||||
Db        120 AAAACCTGGGATGATGCAGAGAGTTTC 146

RESULT 2
US-10-226-420-1
; Sequence 1, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 483
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)...(455)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk2
US-10-226-420-1

Alignment Scores:
Pred. No.:              1.24e-12        Length:         483
Score:                  134.00          Matches:        22
Percent Similarity:     82.76%          Conservative:   2
Best Local Similarity:  75.86%          Mismatches:     5
Query Match:            78.82%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-3 (1-29) x US-10-226-420-1 (1-483)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   ||||||||||||||||||:::   ||||||||||||||||||   :::
Db         60 GATTGTCCCTCTGACTGGTCCTCCTATGATCAGCATTGCTACAAGGTCTTCAGTGAACTC 119

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   ||||||||||||||||
Db        120 AAAACCTGGGATGATGCAGAGAGTTTC 146

RESULT 3
US-09-929-230-4
; Sequence 4, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 721
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)...(546)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk3
US-09-929-230-4

Alignment Scores:
Pred. No.:              2.93e-11        Length:         721
Score:                  127.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            74.71%          Indels:         0
DB:                     10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-929-230-4 (1-721)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||:::   |||||||||:::||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        220 AAGACGTGGGACGATGCAGAGAGGTTC 246

RESULT 4
US-10-226-420-4
; Sequence 4, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 721
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (91)...(546)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk3
US-10-226-420-4

Alignment Scores:
Pred. No.:              2.93e-11        Length:         721
Score:                  127.00          Matches:        21
Percent Similarity:     79.31%          Conservative:   2
Best Local Similarity:  72.41%          Mismatches:     6
Query Match:            74.71%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-3 (1-29) x US-10-226-420-4 (1-721)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||   ||||||||||||:::   |||||||||:::||||||||||||
Db        160 GATTGTCCCTCTGGTTGGTCCTCCTATGATCAGCATTGCTACAGGGTCTTCAAACAACTC 219
```

```
Qy          21 LysThrTrpThrAspAlaGluSerPhe 29
               ||||||||| ||||||||| |||
Db         220 AAGACGTGGGACGATGCAGAGAGGTTC 246

RESULT 5
US-09-929-230-3
; Sequence 3, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
;   NAME/KEY: misc_feature
;   LOCATION: 9, 12, 18, 21, 24, 27, 30, 36, 39, 42, 45, 48, 51, 54, 57,
;   LOCATION: 66, 69, 78, 81, 105, 111, 117, 123, 135, 141, 153, 162,
;   LOCATION: 168, 171, 174, 177, 180, 189, 192, 204, 210, 213, 219, 222,
;   LOCATION: 225, 231, 234, 243, 246, 258, 261, 270, 297, 306, 312, 324
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 330, 333, 336, 339, 348, 351, 357, 360, 366, 372, 390, 393,
;   LOCATION: 405, 408, 411, 417, 432, 435, 438, 441, 447, 450, 453
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-3

Alignment Scores:
Pred. No.:                3.33e-10        Length:          453
Score:                    119.00          Matches:         18
Percent Similarity:       68.97%          Conservative:    2
Best Local Similarity:    62.07%          Mismatches:      9
Query Match:              70.00%          Indels:          0
DB:                       10              Gaps:            0

US-09-938-114-3 (1-29) x US-09-929-230-3 (1-453)

Qy           1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
               ||||||      ||||||      |||:::  ||||||||||||||||||   :::
Db          58 GAYTGYCCNWSNGAYTGGWSNWSNTAYGAYCARCAYTGYTAYAARGTNTTYWSNGARYTN 117

Qy          21 LysThrTrpThrAspAlaGluSerPhe 29
               ||||||||| ||||||||| |||
Db         118 AARACNTGGGAYGAYGCNGARWSNTTY 144

RESULT 6
US-10-226-420-3
; Sequence 3, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 453
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 9, 12, 18, 21, 24, 27, 30, 36, 39, 42, 45, 48, 51, 54, 57,
;   LOCATION: 66, 69, 78, 81, 105, 111, 117, 123, 135, 141, 153, 162,
;   LOCATION: 168, 171, 174, 177, 180, 189, 192, 204, 210, 213, 219, 222,
;   LOCATION: 225, 231, 234, 243, 246, 258, 261, 270, 297, 306, 312, 324
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 330, 333, 336, 339, 348, 351, 357, 360, 366, 372, 390, 393,
;   LOCATION: 405, 408, 411, 417, 432, 435, 438, 441, 447, 450, 453
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-3

Alignment Scores:
Pred. No.:                3.33e-10        Length:          453
Score:                    119.00          Matches:         18
Percent Similarity:       68.97%          Conservative:    2
Best Local Similarity:    62.07%          Mismatches:      9
Query Match:              70.00%          Indels:          0
DB:                       13              Gaps:            0

US-09-938-114-3 (1-29) x US-10-226-420-3 (1-453)

Qy           1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
               ||||||      ||||||      |||:::  ||||||||||||||||||   :::
Db          58 GAYTGYCCNWSNGAYTGGWSNWSNTAYGAYCARCAYTGYTAYAARGTNTTYWSNGARYTN 117

Qy          21 LysThrTrpThrAspAlaGluSerPhe 29
               ||||||||| ||||||||| |||
Db         118 AARACNTGGGAYGAYGCNGARWSNTTY 144

RESULT 7
US-09-929-230-6
; Sequence 6, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
;   LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
;   LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
;   LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-6

Alignment Scores:
Pred. No.:                1.53e-09        Length:          456
Score:                    115.00          Matches:         18
Percent Similarity:       65.52%          Conservative:    1
Best Local Similarity:    62.07%          Mismatches:      10
Query Match:              67.65%          Indels:          0
DB:                       10              Gaps:            0
```

```
US-09-938-114-3 (1-29) x US-09-929-230-6 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||          |||      |||:::   |||||||||   ||||||||||||
Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        130 AARACNTGGGAYGAYGCNGARMGNTTY 156

RESULT 8
US-10-226-420-6
; Sequence 6, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:5.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 9, 21, 24, 30, 33, 36, 39, 42, 48, 51, 54, 57, 60, 63,
;   LOCATION: 66, 69, 78, 81, 84, 90, 93, 114, 117, 129, 135, 147, 153,
;   LOCATION: 162, 171, 177, 180, 186, 189, 192, 201, 204, 213, 216, 222,
;   LOCATION: 225, 231, 234, 237, 246, 249, 252, 255, 261, 276, 279
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 282, 285, 291, 309, 312, 321, 327, 330, 333, 336, 339, 360,
;   LOCATION: 366, 372, 378, 381, 384, 393, 396, 414, 429, 441, 447, 456
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-6

Alignment Scores:
Pred. No.:              1.53e-09        Length:         456
Score:                  115.00          Matches:        18
Percent Similarity:     65.52%          Conservative:   1
Best Local Similarity:  62.07%          Mismatches:     10
Query Match:            67.65%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-3 (1-29) x US-10-226-420-6 (1-456)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||          |||      |||:::   |||||||||   ||||||||||||
Db         70 GAYTGYCCNWSNGGNTGGWSNWSNTAYGAYCARCAYTGYTAYMGNGTNTTYAARCARYTN 129

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              |||||||||   |||||||||   |||
Db        130 AARACNTGGGAYGAYGCNGARMGNTTY 156

RESULT 9
US-09-938-114-1
; Sequence 1, Application US/09938114
; Publication No. US20030022350A1
;   GENERAL INFORMATION:
;        APPLICANT: Chun Wang, Banjamin XY Li, Xin Cheng,
;                   Jing Liu, Li-Wen Niu, Wan-Zhi Huang,
;                   Zhen-Yu Xu, Dan Luo, Lian-Di Kang,
;                   Jin-Guo Ding, Fang Rong, Yan Liu and
;                   Hui-Ran Chen
;        TITLE OF INVENTION: AN ANTITHROMBOSIS ENZYME FROM THE SNAKE
;                            VENOM OF AGKISTRODON ACUTUS
;        NUMBER OF SEQUENCES: 12
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 633 West Fifth Street
;                     Suite 4700
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: U.S.A.
;             ZIP: 90071-2066
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;                          storage
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: IBM P.C. DOS 5.0
;             SOFTWARE: FastSEQ for Windows 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/938,114
;             FILING DATE: 23-Aug-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/058,740
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Chen, Anthony C.
;             REGISTRATION NUMBER: 38,673
;             REFERENCE/DOCKET NUMBER: 233/298
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 544 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-938-114-1

Alignment Scores:
Pred. No.:              1.92e-09        Length:         544
Score:                  115.00          Matches:        20
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            67.65%          Indels:         0
DB:                     11              Gaps:           0

US-09-938-114-3 (1-29) x US-09-938-114-1 (1-544)

Qy         10 GluGlyHisCysTyrLysValPheLysGlnSerLysThrTrpThrAspAlaGluSerPhe 29
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGGGCATTGCTACAAGGTCTTCAAACAATCTAAGACCTGGACAGATGCAGAGAGCTTC 60

RESULT 10
US-09-969-763-2
; Sequence 2, Application US/09969763
; Publication No. US20020198363A1
; GENERAL INFORMATION:
;  APPLICANT: FUKUCHI, NAOYUKI
;  APPLICANT:  KITO, MORIKAZU
;  APPLICANT:  KAYAHARA, TAKASHI
;  APPLICANT:  FUTAKI, FUMIE
;  APPLICANT:  ISHIKAWA, KOHKI
;  APPLICANT:  SUZUKI, EIICHIRO
;  APPLICANT:  GONDOH, KEIKO
;  APPLICANT:  SHIMBA, NOBUHISA
;  APPLICANT:  YAMADA, NAOYUKI
;  TITLE OF INVENTION: PROTEIN HAVING ANTITHROMBOTIC ACTIVITY AND METHOD FOR PRODUCING T
;  FILE REFERENCE: 214760US0
;  CURRENT APPLICATION NUMBER: US/09/969,763
;  CURRENT FILING DATE:  2000-10-25
;  PRIOR APPLICATION NUMBER: JP 2000-305279
```

```
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;    LENGTH: 690
;    TYPE: DNA
;    ORGANISM: Crotalus harridus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (66)..(512)
;    OTHER INFORMATION:
US-09-969-763-2

Alignment Scores:
Pred. No.:                1.18e-08        Length:         690
Score:                    111.00          Matches:        18
Percent Similarity:       72.41%          Conservative:   3
Best Local Similarity:    62.07%          Mismatches:     8
Query Match:              65.29%          Indels:         0
DB:                       10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-969-763-2 (1-690)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||||||||||:::   :::|||||||||   |||||||||
Db        141 GAATGTCCCTCCGGTTGGTCTTCCTATGATCGGTATTGCTACAAGCCCTTCAAACAAGAG 200

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db        201 ATGACCTGGGCCGATGCAGAGAGGTTC 227

RESULT 11
US-09-929-230-7
; Sequence 7, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;    LENGTH: 580
;    TYPE: DNA
;    ORGANISM: Sistrurus miliarius
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (3)...(434)
;    NAME/KEY: misc_feature
;    LOCATION: (0)...(0)
;    OTHER INFORMATION: Zsnk4
US-09-929-230-7

Alignment Scores:
Pred. No.:                1.38e-08        Length:         580
Score:                    110.00          Matches:        17
Percent Similarity:       75.86%          Conservative:   5
Best Local Similarity:    58.62%          Mismatches:     7
Query Match:              64.71%          Indels:         0
DB:                       10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-929-230-7 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||||||||   :::|||:::   :::||||||:::|||   ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              :::||||||   |||||||||   |||
Db         96 AGGACGTGGGACGATGCAGAGAGGTTC 122

RESULT 12
US-10-226-420-7
; Sequence 7, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;    LENGTH: 580
;    TYPE: DNA
;    ORGANISM: Sistrurus miliarius
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (3)...(434)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (0)...(0)
;    OTHER INFORMATION: Zsnk4
US-10-226-420-7

Alignment Scores:
Pred. No.:                1.38e-08        Length:         580
Score:                    110.00          Matches:        17
Percent Similarity:       75.86%          Conservative:   5
Best Local Similarity:    58.62%          Mismatches:     7
Query Match:              64.71%          Indels:         0
DB:                       13              Gaps:           0

US-09-938-114-3 (1-29) x US-10-226-420-7 (1-580)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||   |||||||||   :::|||:::   :::||||||:::|||   ||||||
Db         36 GATTGTCCCTCTGATTGGTATGCCTATGATCAGTATTGCTACAGGGTCATCAAACAACTC 95

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
              :::||||||   |||||||||   |||
Db         96 AGGACGTGGGACGATGCAGAGAGGTTC 122

RESULT 13
US-09-929-230-9
; Sequence 9, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;    LENGTH: 432
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;    OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;    NAME/KEY: misc_feature
;    LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;    LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;    LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;    LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;    OTHER INFORMATION: n = A,T,C or G
;    NAME/KEY: misc_feature
;    LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
```

```
;   OTHER INFORMATION: n = A,T,C or G
US-09-929-230-9

Alignment Scores:
Pred. No.:              1.34e-07        Length:         432
Score:                  103.00          Matches:        16
Percent Similarity:     65.52%          Conservative:   3
Best Local Similarity:  55.17%          Mismatches:     10
Query Match:            60.59%          Indels:         0
DB:                     10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-929-230-9 (1-432)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTY 120

RESULT 14
US-10-226-420-9
; Sequence 9, Application US/10226420
; Publication No. US20030157686A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/10/226,420
;  CURRENT FILING DATE:  2002-08-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 432
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: This degenerate nucleotide sequence encodes the
;   OTHER INFORMATION: amino acid sequence of SEQ ID NO:8.
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 6, 15, 18, 21, 24, 27, 42, 45, 57, 78, 81, 93, 96, 99, 111,
;   LOCATION: 117, 126, 135, 141, 144, 150, 153, 156, 165, 171, 177, 180,
;   LOCATION: 186, 189, 195, 198, 201, 231, 240, 243, 246, 255, 273, 276,
;   LOCATION: 285, 291, 294, 297, 300, 315, 327, 330, 345, 348, 360
;   OTHER INFORMATION: n = A,T,C or G
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 363, 369, 372, 378, 384, 393, 405, 423, 426
;   OTHER INFORMATION: n = A,T,C or G
US-10-226-420-9

Alignment Scores:
Pred. No.:              1.34e-07        Length:         432
Score:                  103.00          Matches:        16
Percent Similarity:     65.52%          Conservative:   3
Best Local Similarity:  55.17%          Mismatches:     10
Query Match:            60.59%          Indels:         0
DB:                     13              Gaps:           0

US-09-938-114-3 (1-29) x US-10-226-420-9 (1-432)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              ||||||      ||||||   :::|||:::   :::||||||   |||   ||||||
Db         34 GAYTGYCCNWSNGAYTGGTAYGCNTAYGAYCARTAYTGYTAYMGNGTNATHAARCARYTN 93

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                ||||||   |||||||||   |||
Db         94 MGNACNTGGGAYGAYGCNGARMGNTTY 120

RESULT 15
US-09-929-230-10
; Sequence 10, Application US/09929230
; Patent No. US20020161203A1
; GENERAL INFORMATION:
;  APPLICANT: Sheppard, Paul O.
;  APPLICANT:  Bishop, Paul D.
;  TITLE OF INVENTION: RATTLESNAKE VENOM GLAND PROTEINS
;  FILE REFERENCE: 00-72
;  CURRENT APPLICATION NUMBER: US/09/929,230
;  CURRENT FILING DATE:  2001-08-13
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Sistrurus miliarius
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (88)...(561)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Zsnk5
US-09-929-230-10

Alignment Scores:
Pred. No.:              2.52e-06        Length:         725
Score:                  97.00           Matches:        15
Percent Similarity:     72.41%          Conservative:   6
Best Local Similarity:  51.72%          Mismatches:     8
Query Match:            57.06%          Indels:         0
DB:                     10              Gaps:           0

US-09-938-114-3 (1-29) x US-09-929-230-10 (1-725)

Qy          1 AspCysSerSerAspTrpSerSerTyrGluGlyHisCysTyrLysValPheLysGlnSer 20
              :::|||   |||   |||   :::|||:::   :::||||||:::|||   |||:::
Db        163 AATTGTCCCTCTGGTTGGTTCGCCTACGATCAGTATTGCTACAGGGTCATCAAACGACTC 222

Qy         21 LysThrTrpThrAspAlaGluSerPhe 29
                |||||||||   |||||||||   |||
Db        223 AAGACCTGGGACGATGCAGAGCGGTTC 249


Search completed: December  8, 2003, 19:36:05
Job time : 116.337 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:49:54 ; Search time 5.8 Seconds
(without alignments)
480.843 Million cell updates/sec

Title: US-09-938-114-3
Perfect score: 170
Sequence: 1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	158	92.9	152	2	JC7134	agkisacutacin alph
2	133	78.2	146	2	JC4691	coagulation factor
3	132	77.6	146	2	JC7135	agkisacutacin beta
4	129	75.9	125	2	JC5059	bitiscetin beta ch
5	128	75.3	129	2	JC4329	coagulation factor
6	124	72.9	146	2	JC7105	aggretin beta chai
7	123	72.4	152	2	JC4690	coagulation factor
8	122	71.8	40	2	S56007	tokaracetin beta c
9	122	71.8	131	2	JC5058	bitiscetin alpha c
10	119	70.0	125	2	B47267	botrocetin beta ch
11	117	68.8	133	2	A47267	botrocetin alpha c
12	115	67.6	40	2	B56829	alboaggregin-B alp
13	114	67.1	123	2	JC2415	echicetin beta cha
14	113	66.5	40	2	S56006	tokaracetin alpha
15	112	65.9	30	2	A53088	factor IX/factor X
16	112	65.9	123	2	B42972	coagulation factor
17	104	61.2	40	2	A56829	alboaggregin-B bet
18	103	60.6	144	2	PC7027	aggretin alpha cha
19	100.5	59.1	29	2	PC4421	multactivase (EC 3
20	97	57.1	30	2	B53088	factor IX/factor X
21	96	56.5	135	2	A38609	lectin, galactose-
22	95	55.9	2109	1	I50421	aggrecan precursor
23	94	55.3	2415	1	A39086	aggrecan precursor
24	92	54.1	1340	2	A39808	proteoglycan core
25	92	54.1	2327	2	T42630	aggrecan - bovine
26	90	52.9	2124	2	A28452	proteoglycan core
27	90	52.9	2132	1	A55182	aggrecan precursor
28	89	52.4	71	2	S55679	flavocetin A - hab
29	84	49.4	1257	2	S28764	neurocan precursor
30	84	49.4	1268	2	S52781	neurocan - mouse
31	83	48.8	1458	1	A49707	phospholipase A2 r
32	82	48.2	883	2	S57653	brevican precursor
33	80	47.1	175	2	A37194	pancreatic thread
34	80	47.1	883	2	S49126	brevican precursor
35	79	46.5	912	2	A54423	brevican precursor
36	76	44.7	223	2	B46467	NKR-P1 protein hom
37	76	44.7	330	2	T46256	brevican - human (
38	76	44.7	3562	2	A47171	chondroitin sulfat
39	75	44.1	147	2	A26697	echinoidin - sea u
40	75	44.1	862	2	S43922	versican - pig-tai
41	75	44.1	1326	2	B56395	secretory phosphol
42	75	44.1	1465	2	A56395	secretory phosphol
43	75	44.1	1643	2	T14274	versican precursor
44	75	44.1	2397	1	A55535	versican precursor
45	75	44.1	2409	1	A60979	versican precursor

ALIGNMENTS

RESULT 1
JC7134
agkisacutacin alpha chain precursor - sharp-nosed viper
N;Alternate names: fibrinogenlytic venom protein
C;Species: Agkistrodon acutus (sharp-nosed viper)
C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000
C;Accession: JC7134; PC7037
R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.
Biochem. Biophys. Res. Commun. 265, 530-535, 1999
A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenlytic venom
A;Reference number: JC7134; MUID:20025379; PMID:10558903
A;Accession: JC7134
A;Molecule type: mRNA
A;Residues: 1-152 <CHE>
A;Cross-references: GB:AF176420
A;Experimental source: venom gland
A;Accession: PC7037
A;Molecule type: protein
A;Residues: 24-53;84-86;87-94;125-136;137-152 <CH2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: agkisacutacin alpha chain #status experimental <MAT>

Query Match 92.9%; Score 158; DB 2; Length 152;
Best Local Similarity 93.1%; Pred. No. 7.9e-13;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |||| ||||||||||||||||| |||||
Db         24 DCSSGWSSYEGHCYKVFKQSKTWADAESF 52
```

RESULT 2
JC4691
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4691; B39332; JC4330
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4691
A;Molecule type: mRNA
A;Residues: 1-146 <MAT1>
A;Cross-references: DDBJ:D83332; NID:g1402641; PIDN:BAA11888.1; PID:g1402642
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
otein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.

```
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: B39332
A;Molecule type: protein
A;Residues: 24-146 <ATO>
R;Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 118, 965-973, 1995
A;Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flav
A;Reference number: JC4329; MUID:96318509; PMID:8749314
A;Accession: JC4330
A;Molecule type: protein
A;Residues: 24-146 <AT2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; lectin; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-146/Product: factor IX/X binding protein chain B #status predicted <MAT>
F;25-142/Domain: C-type lectin homology <LCH>
F;25-36,53-142,119-134/Disulfide bonds: #status predicted

  Query Match             78.2%;  Score 133;  DB 2;  Length 146;
  Best Local Similarity   72.4%;  Pred. No. 1.1e-09;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||||||||| | : | | |||:|
Db         24 DCPSDWSSYEGHCYKPFSEPKNWADAENF 52


RESULT 3
JC7135
agkisacutacin beta chain precursor - sharp-nosed viper
N;Alternate names: fibrinogenlytic venom protein
C;Species: Agkistrodon acutus (sharp-nosed viper)
C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 24-Oct-2000
C;Accession: JC7135; PC7038
R;Cheng, X.; Qian, Y.; Liu, Q.; Li, B.X.Y.; Zhang, M.; Liu, J.
Biochem. Biophys. Res. Commun. 265, 530-535, 1999
A;Title: Purification, characterization, and cDNA cloning of a new fibrinogenlytic venom
A;Reference number: JC7134; MUID:20025379; PMID:10558903
A;Accession: JC7135
A;Molecule type: mRNA
A;Residues: 1-146 <CHE>
A;Cross-references: GB:AF176421
A;Experimental source: venom gland
A;Accession: PC7038
A;Molecule type: protein
A;Residues: 24-50;59-83;102-107;112-114 <CH2>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer; venom
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-146/Product: agkisacutacin beta chain #status experimental <MAT>

  Query Match             77.6%;  Score 132;  DB 2;  Length 146;
  Best Local Similarity   72.4%;  Pred. No. 1.4e-09;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |:|||||||||| | : ||| ||| |
Db         24 DCPSEWSSYEGHCYKPFDEPKTWADAEKF 52


RESULT 4
JC5059
bitiscetin beta chain - puff adder
C;Species: Bitis arietans (puff adder)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 12-Feb-1999
C;Accession: JC5059; JC5917
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
submitted to JIPID, January 1997
A;Description: Complete amino acid sequence of bitiscetin, a novel von willebrand factor
A;Reference number: JC5058
A;Accession: JC5059
A;Molecule type: protein
A;Residues: 1-125 <MAT>
A;Experimental source: snake venom
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oze
Res. Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997
A;Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor modula
A;Reference number: JC5916
A;Accession: JC5917
A;Molecule type: protein
A;Residues: 1-125 <MA2>
A;Experimental source: venom
C;Comment: This protein is a modulator of a von Willebrand factor modulator.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: venom
F;4-121/Domain: C-type lectin homology <LCH>

  Query Match             75.9%;  Score 129;  DB 2;  Length 125;
  Best Local Similarity   75.0%;  Pred. No. 3e-09;
  Matches   21;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              |  |||||:||||||||  ||| ||| |
Db          4 CLPDWSSYKGHCYKVFKVEKTWADAEKF 31


RESULT 5
JC4329
coagulation factor IX-binding protein A chain - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 06-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 13-Mar-1998
C;Accession: JC4329
R;Atoda, H.; Ishikawa, M.; Yoshihara, E.; Sekiya, F.; Morita, T.
J. Biochem. 118, 965-973, 1995
A;Title: Blood coagulation factor IX-binding protein from the venom of Trimeresurus flavo
A;Reference number: JC4329; MUID:96318509; PMID:8749314
A;Accession: JC4329
A;Molecule type: protein
A;Residues: 1-129 <ATO>
C;Comment: This protein binds calcium.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; calcium binding; venom
F;2-127/Domain: C-type lectin homology <LCH>
F;2-13,30-127,102-119/Disulfide bonds: #status predicted

  Query Match             75.3%;  Score 128;  DB 2;  Length 129;
  Best Local Similarity   75.9%;  Pred. No. 4.1e-09;
  Matches   22;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | |||||||||| ||  ||| ||| |
Db          1 DCPSGWSSYEGHCYKPFKLYKTWDDAERF 29


RESULT 6
JC7105
aggretin beta chain - Malayan pit viper
C;Species: Calloselasma rhodostoma (Malayan pit viper)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7105
R;Chung, C.H.; Au, L.C.; Huang, T.F.
Biochem. Biophys. Res. Commun. 263, 723-727, 1999
A;Title: Molecular cloning and sequence analysis of aggretin, a collagen-like platelet ag
A;Reference number: PC7027; MUID:99443731; PMID:10512747
A;Accession: JC7105
A;Molecule type: mRNA
A;Residues: 1-146 <CHU>
A;Experimental source: venom gland
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; platelet aggregation; venom

  Query Match             72.9%;  Score 124;  DB 2;  Length 146;
  Best Local Similarity   69.0%;  Pred. No. 1.4e-08;
  Matches   20;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
```

```
                ||  |  |||||||||||  |  :  |  |  |||  |
Db          24 DCPSGWSSYEGHCYKPFNEPKNWADAERF 52


RESULT 7
JC4690
coagulation factor IX/factor X-binding protein chain A precursor - habu
C;Species: Trimeresurus flavoviridis (habu)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 20-Jun-2000
C;Accession: JC4690; A39332
R;Matsuzaki, R.; Yoshiara, E.; Yamada, M.; Shima, K.; Atoda, H.; Morita, T.
Biochem. Biophys. Res. Commun. 220, 382-387, 1996
A;Title: cDNA cloning of IX/X-BP, a heterogeneous two-chain anticoagulant protein from s
A;Reference number: JC4690; MUID:96184662; PMID:8645314
A;Accession: JC4690
A;Molecule type: mRNA
A;Residues: 1-152 <MAT1>
A;Cross-references: DDBJ:D83331; NID:g1402639; PIDN:BAA11887.1; PID:g1402640
A;Experimental source: venom
R;Atoda, H.; Hyuga, M.; Morita, T.
J. Biol. Chem. 266, 14903-14911, 1991
A;Title: The primary structure of coagulation factor IX/factor X-binding protein isolate
otein, tetranectin, and lymphocyte Fc epsilon receptor for immunoglobulin E.
A;Reference number: A39332; MUID:91332000; PMID:1831197
A;Accession: A39332
A;Status: preliminary
A;Molecule type: protein
A;Residues: 24-152 <ATO>
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; blood coagulation; glycoprotein; hemolymph; lectin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-152/Product: factor IX/X binding protein chain A #status predicted <MAT>
F;25-150/Domain: C-type lectin homology <LCH>
F;25-36,53-150,125-142/Disulfide bonds: #status predicted

  Query Match             72.4%;  Score 123;  DB 2;  Length 152;
  Best Local Similarity   74.1%;  Pred. No. 2e-08;
  Matches   20;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAE 27
                ||  |  ||||||||||  |::  |||  |||
Db         24 DCLSGWSSYEGHCYKAFEKYKTWEDAE 50


RESULT 8
S56007
tokaracetin beta chain - Trimeresurus tokarensis (fragment)
N;Alternate names: platelet aggregation inhibitor; platelet antagonist
C;Species: Trimeresurus tokarensis
C;Date: 10-Oct-1995 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
C;Accession: S56007
R;Kawasaki, T.; Taniuchi, Y.; Hisamichi, N.; Fujimura, Y.; Suzuki, M.; Titani, K.; Sakai
Biochem. J. 308, 947-953, 1995
A;Title: Tokaracetin, a new platelet antagonist that binds to platelet glycoprotein Ib a
A;Reference number: S56006; MUID:97104297; PMID:8948455
A;Accession: S56007
A;Molecule type: protein
A;Residues: 1-40 <KAW>

  Query Match             71.8%;  Score 122;  DB 2;  Length 40;
  Best Local Similarity   65.5%;  Pred. No. 8.4e-09;
  Matches   19;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
                ||  ||||||:  |||:||:|    |  |||  |
Db          1 DCPSDWSSYDEHCYRVFQQKMNWEDAEKF 29


RESULT 9
JC5058
bitiscetin alpha chain - puff adder
N;Alternate names: von Willebrand factor modulator protein
C;Species: Bitis arietans (puff adder)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 12-Feb-1999
C;Accession: JC5058; JC5916
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
submitted to JIPID, January 1997
A;Description: Complete amino acid sequence of bitiscetin, a novel von willebrand factor
A;Reference number: JC5058
A;Contents: snake venom
A;Accession: JC5058
A;Molecule type: protein
A;Residues: 1-131 <MAT>
A;Experimental source: snake venom
R;Matsui, T.; Hamako, J.; Suzuki, M.; Hayashi, N.; Ito, M.; Makita, K.; Fujimura, Y.; Oz
Res. Commun. Biochem. Cell Mol. Biol. 1, 271-284, 1997
A;Title: Complete amino acid sequence of bitiscetin, a novel von Willebrand factor modul
A;Reference number: JC5916
A;Accession: JC5916
A;Molecule type: protein
A;Residues: 1-131 <MA2>
A;Experimental source: venom
C;Comment: This protein is a modulator of a von Willebrand factor modulator.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: venom
F;4-125/Domain: C-type lectin homology <LCH>

  Query Match             71.8%;  Score 122;  DB 2;  Length 131;
  Best Local Similarity   71.4%;  Pred. No. 2.3e-08;
  Matches   20;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 CSSDWSSYEGHCYKVFKQSKTWTDAESF 29
                |  |||||:||||||||:  ||  |||  |
Db          4 CLPDWSSYKGHCYKVFKKVGTWEDAEKF 31


RESULT 10
B47267
botrocetin beta chain - jararaca
N;Alternate names: two chain botrocetin beta chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: B47267; C37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator pu
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: B47267
A;Molecule type: protein
A;Residues: 1-125 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124086)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: C37958
A;Molecule type: protein
A;Residues: 1-40 <FUJ>
C;Complex: heterodimer of alpha (see PIR:A47267) and beta chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-121/Domain: C-type lectin homology <LCH>
F;2-13,30-121,98-113/Disulfide bonds: #status experimental
F;75/Disulfide bonds: interchain (to alpha-80) #status experimental

  Query Match             70.0%;  Score 119;  DB 2;  Length 125;
  Best Local Similarity   65.5%;  Pred. No. 5.4e-08;
  Matches   19;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
                ||  ||||||||||:  ||:    |  |||  |
Db          1 DCPPDWSSYEGHCYRFFKEWMHWDDAEEF 29


RESULT 11
```

```
A47267
botrocetin alpha chain - jararaca
N;Alternate names: two chain botrocetin alpha chain
C;Species: Bothrops jararaca (jararaca)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 15-May-1998
C;Accession: A47267; B37958
R;Usami, Y.; Fujimura, Y.; Suzuki, M.; Ozeki, Y.; Nishio, K.; Fukui, H.; Titani, K.
Proc. Natl. Acad. Sci. U.S.A. 90, 928-932, 1993
A;Title: Primary structure of two-chain botrocetin, a von Willebrand factor modulator pu
A;Reference number: A47267; MUID:93157385; PMID:8430107
A;Accession: A47267
A;Molecule type: protein
A;Residues: 1-133 <USA>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:124085)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: B37958
A;Molecule type: protein
A;Residues: 1-40 <FUJ>
C;Complex: heterodimer of alpha and beta (see PIR:B47267) chains
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: hemagglutinin; heterodimer; venom
F;2-128/Domain: C-type lectin homology <LCH>
F;2-13,30-128,103-120/Disulfide bonds: #status experimental
F;80/Disulfide bonds: interchain (to beta-75) #status experimental

  Query Match             68.8%;  Score 117;  DB 2;  Length 133;
  Best Local Similarity   65.5%;  Pred. No. 1e-07;
  Matches   19;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | |||||:||| |:|   | ||| |
Db          1 DCPSGWSSYEGNCYKFFQQKMNWADAERF 29


RESULT 12
B56829
alboaggregin-B alpha chain - green pit viper (fragment)
C;Species: Trimeresurus albolabris (green pit viper)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 03-May-1996
C;Accession: B56829
R;Yoshida, E.; Fujimura, Y.; Miura, S.; Sugimoto, M.; Fukui, H.; Narita, N.; Usami, Y.;
Biochem. Biophys. Res. Commun. 191, 1386-1392, 1993
A;Title: Alboaggregin-B and botrocetin, two snake venom proteins with highly homologous
A;Reference number: A56829; MUID:93221514; PMID:8466514
A;Accession: B56829
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-40 <YOS>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:128751)
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: disulfide bond; heterodimer

  Query Match             67.6%;  Score 115;  DB 2;  Length 40;
  Best Local Similarity   62.1%;  Pred. No. 6.4e-08;
  Matches   18;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||||:: :||:: || ||| ||| |
Db          1 DCPSDWSSFKQYCYQIIKQLKTWEDAERF 29


RESULT 13
JC2415
echicetin beta chain - saw-scaled viper
C;Species: Echis carinatus (saw-scaled viper)
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 15-Oct-1996
C;Accession: JC2415
R;Peng, M.; Holt, J.C.; Niewiarowski, S.
Biochem. Biophys. Res. Commun. 205, 68-72, 1994
A;Title: Isolation, characterization and amino acid sequence of echicetin beta subunit,
A;Reference number: JC2415; MUID:95091801; PMID:7999097
A;Accession: JC2415
A;Molecule type: protein
A;Residues: 1-123 <PEN>
A;Experimental source: venom
C;Comment: This protein inhibits agglutination of fixed platelets induced by several plat
llebrand factor and alboaggregins.
C;Superfamily: tetranectin; C-type lectin homology
C;Keywords: anticoagulant; dimer
F;2-119/Domain: C-type lectin homology <LCH>
F;2-13,30-119,96-111/Disulfide bonds: #status predicted

  Query Match             67.1%;  Score 114;  DB 2;  Length 123;
  Best Local Similarity   62.1%;  Pred. No. 2.3e-07;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              :|  ||| |||:||||||:   | ||| |
Db          1 NCLPDWSVYEGYCYKVFKERMNWADAEKF 29


RESULT 14
S56006
tokaracetin alpha chain - Trimeresurus tokarensis (fragment)
N;Alternate names: platelet aggregation inhibitor; platelet antagonist
C;Species: Trimeresurus tokarensis
C;Date: 10-Oct-1995 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
C;Accession: S56006
R;Kawasaki, T.; Taniuchi, Y.; Hisamichi, N.; Fujimura, Y.; Suzuki, M.; Titani, K.; Sakai,
Biochem. J. 308, 947-953, 1995
A;Title: Tokaracetin, a new platelet antagonist that binds to platelet glycoprotein Ib ar
A;Reference number: S56006; MUID:97104297; PMID:8948455
A;Accession: S56006
A;Molecule type: protein
A;Residues: 1-40 <KAW>
C;Superfamily: tetranectin; C-type lectin homology

  Query Match             66.5%;  Score 113;  DB 2;  Length 40;
  Best Local Similarity   65.5%;  Pred. No. 1.1e-07;
  Matches   19;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || | |||:: :||| ||| ||| ||| |
Db          1 DCPSGWSSFKQYCYKPFKQLKTWEDAERF 29


RESULT 15
A53088
factor IX/factor X-binding anticoagulant protein B chain - jararaca (fragment)
C;Species: Bothrops jararaca (jararaca)
C;Date: 13-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996
C;Accession: A53088
R;Sekiya, F.; Atoda, H.; Morita, T.
Biochemistry 32, 6892-6897, 1993
A;Title: Isolation and characterization of an anticoagulant protein homologous to botroce
A;Reference number: A53088; MUID:93326575; PMID:8334120
A;Accession: A53088
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-30 <SEK>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:135337)

  Query Match             65.9%;  Score 112;  DB 2;  Length 30;
  Best Local Similarity   62.1%;  Pred. No. 1.2e-07;
  Matches   18;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 DCSSDWSSYEGHCYKVFKQSKTWTDAESF 29
              || |||| ||||||:|| : :   ||| |
Db          1 DCPSDWSPYEGHCYRVFTEPQNXADAEKF 29
```